
Modellers as influencers?
J.O.E. Remmers
Modellers as influencers?
analysing practices and standards in hydrodynamic decision-support modelling
J.O.E. Remmers

Propositions

1 | Transparent and retraceable modelling practice requires consistent inconsistency.
(this thesis)

2 | Reflexivity and critical thinking have to be integrated in the hydrological modelling curriculum.
(this thesis)

3 | To write right propositions is to be wrong.

4 | The positivist scientific language inhibits discussion about subjectivity.

5 | Doing a PhD resembles the story of Pandora’s box.

6 | To ‘Act normal’ is to act without creativity.

7 | Humans as creatures of habit prevent conscious change.

Propositions belonging to the thesis, entitled

Modellers as influencers? Analysing practices and standards in hydrodynamic decision-support modelling

J.O.E. Remmers
Wageningen, 13 February 2025

Modellers as influencers?

Analysing practices and standards in hydrodynamic decision-support modelling

Janneke Oda Elisabeth Remmers

Thesis committee

Promotor
Dr A. J. Teuling
Associate professor, Hydrology and Environmental Hydraulics
Wageningen University & Research

Co-promotor
Dr L. A. Melsen
Associate professor, Hydrology and Environmental Hydraulics
Wageningen University & Research

Other members
Prof. Dr G. J. Hofstede, Wageningen University & Research
Prof. Dr T. Krüger, Humboldt University, Berlin, Germany
Prof. Dr Mr N. Doorn, TU Delft, Delft, The Netherlands
Prof. Dr M. Z. Zwarteveen, IHE, Delft, The Netherlands

This research was conducted under the auspices of the Graduate School for Socio-Economic and Natural Sciences of the Environment (SENSE).

Modellers as influencers?

Analysing practices and standards in hydrodynamic decision-support modelling

Janneke Oda Elisabeth Remmers

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr C. Kroeze
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Thursday 13 February 2025
at 10.30 am in the Omnia Auditorium.

J. O. E. Remmers
Modellers as influencers? - Analysing practices and standards in hydrodynamic decision-support modelling
xiii+196 pages.

PhD thesis, Wageningen University, Wageningen, the Netherlands (2025)
With references, with summaries in English and Dutch

DOI 10.18174/681291

Summary

KNOWLEDGE and data are cornerstones of many current-day societies in the world. Computer models, as a way to create knowledge, are used throughout science and practice, for many applications and within many disciplines. Throughout the modelling process, uncertainty is introduced, which might lead to differences in results. A factor introducing differences in the model results is the variation in modelling practices, as choosing a method introduces differences in results. To be able to put the results in perspective, it is important to understand how decisions within the modelling process are made. As a modeller makes these modelling decisions based on their own context, inter-modeller variability occurs. Standardisation is one potential way to counter the inter-modeller variability. In water management, several standards have been explored. **Chapter 1** delineates the context and research problem of this thesis.

In the first three research chapters, modelling practices and standards are analysed, informed by interviews with modellers. In **Chapter 2**, I discuss how modellers motivate their modelling decisions. These results are based on fourteen interviews conducted with Dutch hydrodynamic decision-support modellers, either working at local water authorities or at consulting companies. **Chapter 3** relates how modellers perceive automation as a standardisation approach. Automation as a standardisation approach means replacing manual modelling decisions with a computer programming script that executes those decisions without human intervention. Centralised automation has the potential to ensure consistency in modelling decisions. **Chapter 4** contains an analysis of how modellers perceive two different standardisation approaches. I address two case studies: a top-down standardisation approach currently implemented in Australia and a bottom-up standardisation approach that exists in the Netherlands. Top-down and bottom-up refer to how the standardisation approach is initiated. With a top-down initiation, the approach starts at, for instance, a governmental level, while with a bottom-up initiation at the modeller's level.

First, the motivations behind modelling decisions are analysed in **Chapter 2**. I executed an inductive content analysis on the interview transcripts, which led to a classification of eight motivation categories: Individual, Team, Organisational, External, Commissioner, National, International and Consequential. Furthermore, two overarching themes are identified: Vision and Standards. Additionally, the eight motivation categories are compared within different modelling steps. The individual modellers' influence mainly extends to the model implementation decisions. Other decisions, such as model software, are mostly determined by Organisational, External, Commissioner and National considerations. Moreover, in general, feasibility appears to be a more important motivation than reliability and usefulness. The findings indicate that the modeller's sphere of influence is restricted by other factors, such as Organisational or National considerations.

Second, I have studied how modellers perceive automation as a standardisation approach (**Chapter 3**). I used an automation initiative in the Netherlands as my case study. This initiative was a collaboration between research institutes, consulting companies and water authorities. I analysed a different part of the transcripts that I also used in **Chapter 2**. The analysis was based on deductive and inductive content analyses. The analysis has shown that automation has the potential to improve consistency, efficiency and transparency when the modelling process is automated at the organisational or inter-organisational level. However, this does need several requirements to be met: good documentation, clear ownership, adequate maintenance and frequent evaluation. The interviews have also indicated that disadvantages of automation are loss of flexibility, agency of the modeller and insight in uncertainty. Balancing the risks and advantages of automation needs careful consideration of the power between modellers and programmers.

Third, I have explored how modellers perceive bottom-up and top-down standardisation approaches (**Chapter 4**). In this chapter, I compare two cases: a top-down common modelling framework currently implemented in Australia and a bottom-up modelling guideline that exists in the Netherlands. I conducted twenty interviews in Australia and sixteen in the Netherlands with users and developers of the standardisation approaches. The transcripts were analysed with an inductive content analysis. From the interviews, it appeared that consensus is necessary for a standardisation approach to be implemented. However, this is not a guarantee for a successful and widely supported implementation of the standardisation approach. It can be promoted by a top-down approach or external factors, such as an organisation or commissioner promoting it. The interviewees value the advantages standardisation can offer, such as improved transparency and consistency. However, a standardisation approach does need enough resources, maintenance and adaptability to be able to evolve, similar to what is found in **Chapter 3**. Moreover, standardisation has several potential pitfalls: reduced flexibility, assuming one approach is the most desirable, and potentially marginalising certain stakeholders or concepts. As such, standardisation approaches are non-neutral tools, requiring deliberate development, application, evolution and maintenance.

Finally, I have combined my findings of **Chapters 2 - 4** to give recommendations for moving forward in the hydrological modelling network in **Chapter 5**. These recom-

mendations are based on my own and co-author experiences and insights from the critical social sciences. The main take-away, from my perspective, is that responsible modelling is a shared responsibility. I realise that modellers tend to already bear a lot of the responsibility and are the easiest ones to ask actions from. Therefore, I target the full modelling network – from commissioner, to modeller, to end-user – with my recommendations.

In **Chapter 6**, I synthesise my findings and insights. My thesis highlights the social aspects prevalent in hydrodynamic decision-support modelling. Although the responsibility of and differences in modelling results are often ascribed to the individual modeller, this thesis shows that the context in which a modeller works is just as, if not more, important in influencing the model results. As such, in summary, modellers have their own sphere of influence in the modelling process and outside of it, but this is limited and influenced by the modellers' context. One could even say that modellers are influencees – the ones being influenced, since society and the modelling network shape the modellers, maybe even more than the modeller shapes the model.

Samenvatting

KENNIS en data zijn hoekstenen van vele hedendaagse maatschappijen in de hele wereld. Rekenmodellen, als een manier om kennis te creëren, worden in de wetenschap en praktijk veelvuldig en binnen veel disciplines toegepast. Onzekerheden die geïntroduceerd worden tijdens het modelleerprocess kunnen voor verschillen in model resultaten zorgen. Een bepalende factor voor variaties in de modelresultaten is het verschil in modelleerpraktijken, omdat de keuze van een methode verschillen in resultaten introduceert. Om de resultaten in perspectief te kunnen plaatsen, is het belangrijk om te begrijpen hoe beslissingen binnen het modelleerproces worden genomen. Een modelleur neemt immers deze modelleerbeslissingen op basis van hun eigen context, waardoor inter-modelleurvariabiliteit optreedt. Standaardisatie is een mogelijke manier om de inter-modelleurvariabiliteit tegen te gaan. In het waterbeheer zijn verschillende standaarden onderzocht. In **Hoofdstuk 1** belicht ik de context van mijn onderzoek en mijn onderzoeksvragen.

In de eerste drie onderzoekshoofdstukken (**Hoofdstukken 2 - 4**) worden modelleerpraktijken en -standaarden geanalyseerd. In **Hoofdstuk 2** wordt besproken hoe modelleurs hun modelleerbeslissingen motiveren. Deze resultaten zijn gebaseerd op veertien interviews die zijn uitgevoerd met Nederlandse, hydrodynamische, besluitvormingsondersteunende modelleurs, die werkzaam zijn bij lokale waterschappen of adviesbureaus. **Hoofdstuk 3** beschrijft hoe modelleurs automatisering als standaardisatiebenadering zien. Dit hoofdstuk is gebaseerd op de transcripten die zijn verkregen uit de interviews met de Nederlandse hydrodynamische besluitvormingsondersteunende modelleurs. **Hoofdstuk 4** bevat een analyse van hoe modelleurs twee verschillende standaardisatiebenaderingen zien. Ik behandel twee *case studies*: een *top-down* standaardisatiebenadering in Australië en een *bottom-up* standaardisatiebenadering in Nederland. *Top-down* en *bottom-up* refereren naar hoe de standaardisatiebenadering begint. Bij een *top-down* begin wordt de benadering bijvoorbeeld geïnitieerd door een overheid, terwijl bij een *bottom-up* initiatief dit gebeurt door bijvoorbeeld modelleurs zelf.

Als eerste analyseerde ik de motivaties achter modelleerbeslissingen in **Hoofdstuk 2**. Ik voerde een inductieve inhoudsanalyse uit op de interviewtranscripten, wat heeft geleid tot een classificatie van acht motivatiecategorieën: Individual, Team, Organisational, External, Commissioner, National, International en Consequential. Verder werden twee overkoepelende thema's geïdentificeerd: Visie en Normen. Daarnaast werden de acht motivatiecategorieën vergeleken tussen verschillende modelleerstappen. De invloed van de individuele modelleur strekt zich voornamelijk uit tot de beslissin-

gen over de implementatie van het model. Andere beslissingen, zoals modelsoftware, worden meestal bepaald door organisatorische, externe, en nationale overwegingen, als ook door argumenten vanuit de opdrachtgever. Bovendien blijkt haalbaarheid in het algemeen een belangrijkere motivatie te zijn dan betrouwbaarheid en bruikbaarheid. De bevindingen geven aan dat de invloedssfeer van de modelleur wordt beperkt door andere factoren, zoals organisatorische of nationale overwegingen.

Ten tweede heb ik onderzocht hoe modelleurs automatisering zien als een standaardisatiemethode (**Hoofdstuk 3**). Ik heb een automatiseringsinitiatief in Nederland gebruikt als mijn *case study*. Dit initiatief was een samenwerking tussen onderzoeksinstituten, adviesbureaus en waterschappen. Ik heb een ander deel van de transcripten geanalyseerd die ik ook heb gebruikt in **Hoofdstuk 2**. De analyse is gebaseerd op deductieve en inductieve inhoudsanalyses. De analyse heeft aangetoond dat automatisering de consistentie, efficiëntie en transparantie kan verbeteren wanneer het modelleerproces op organisatorisch of interorganisationeel niveau wordt geautomatiseerd. Hiervoor moet echter wel aan een aantal vereisten worden voldaan: goede documentatie, duidelijk eigenaarschap, adequaat onderhoud en frequente evaluatie. De geïnterviewden gaven nadelen van deze automatisering aan: verlies aan flexibiliteit en handelingsvrijheid van de modelleur en inzicht in onzekerheid waren. Om evenwicht te brengen tussen de risico's en voordelen van automatisering vereist een zorgvuldige afweging van de invloed tussen modelleurs en programmeurs.

Ten derde heb ik onderzocht hoe modelleurs *bottom-up* en *top-down* standardisatie methodes ervaren (**Hoofdstuk 4**). In dit hoofdstuk vergelijk ik twee *case studies*: een gemeenschappelijk modelleerraamwerk in Australië en een modelleerhandboek in Nederland. Ik heb twintig interviews in Australië en zestien interviews in Nederland gehouden. De transcripten heb ik geanalyseerd met een inductieve inhoudelijke analyse, waaruit blijkt dat consensus nodig om een standardisatie methode te implementeren. Maar dit is echter geen garantie dat het breed gedragen en succesvol geïmplementeerd wordt. Implementatie kan bevorderd worden door *top-down* of externe druk, zoals organisaties of opdrachtgevers die het gebruik van de standaardisatiebenadering stimuleren. De geïnterviewden waarderen de voordelen die standardisatie met zich mee brengt, bijvoorbeeld consistentie. Maar, een standaardisatie methode heeft genoeg middelen, onderhoud en aanpassingsmogelijkheden nodig om zich door de tijd te kunnen ontwikkelen. Dit is vergelijkbaar met de bevindingen van **Hoofdstuk 3**. Standaardisatie heeft ook meerdere valkuilen: minder flexibiliteit, één perspectief wordt gekozen als voorkeur, en mogelijk worden bepaalde belanghebbenden of concepten gemarginaliseerd. Daarom zijn standaardisatie methodes niet-neutrale hulpmiddelen, waardoor het nodig is om die zorgvuldig te ontwikkelen, te implementeren, te evolueren en te onderhouden.

Ten vierde heb ik de bevindingen van **Hoofdstukken 2 - 4** gecombineerd om aanbevelingen te maken voor de hydrologische modelleergemeenschap in **Hoofdstuk 5**. Deze aanbevelingen zijn gebaseerd op mijn eigen en co-auteur ervaringen en inzichten van de kritische sociale wetenschappen. De hoofdboodschap, vanuit mijn perspectief, is dat verantwoordelijk modelleren een gedeelde verantwoordelijkheid is. Ik realiseer me dat modelleurs al vaak veel verantwoordelijkheid dragen en tegelijkertijd degenen zijn op wie gemakkelijk een beroep wordt gedaan. Daarom richten mijn aanbevelingen zich op de volledige modelleergemeenschap – van opdrachtgever tot modelleur tot eindgebruiker.

In **Hoofdstuk 6** bediscussieer ik mijn bevindingen en inzichten. Mijn thesis benadrukt de sociale aspecten aanwezig in hydrodynamisch modelleren die besluitvorming ondersteunen. Ondanks dat de verantwoordelijkheid voor en de verschillen in modelresultaten vaak worden toegeschreven aan de individuele modelleur, laat deze thesis zien dat de context waarin een modelleur werkt net zo belangrijk, of zelfs belangrijker, is in het beïnvloeden van de model resultaten. Kortom, modelleurs hebben hun eigen invloedssfeer tijdens het modelleerproces en daarbuiten, maar deze wordt gelimiteerd en beïnvloed door de context van de modelleur. Ook de modelleur zelf kan gezien worden als onderhevig aan invloeden, aangezien de maatschappij en de modelleergemeenschap hun sporen op de modelleur nalaten, misschien zelfs meer dan dat de modelleur het model beïnvloedt.

Contents

1

Introduction

"I did not know before that you were a studier of character.
It must be an amusing study."

—Jane Austen, *Pride and Prejudice* (1813)

1.1 | Knowledge creation and models

Knowledge and data are cornerstones in many current-day societies around the world. For example, many rely on the latest traffic news or weather forecasts to plan their journey home after a day of work (de Campos Vallim & Akabane, 2023; Lazo et al., 2009). This is only possible due to knowledge creation – data collection, process understanding and result interpretation (Bereiter & Scardamalia, 2014; Nonaka & Toyama, 2003). Modelling is one form of knowledge creation (Knutti et al., 2013; Savenije, 2009; Schmolke et al., 2010). In weather forecasts, data are collected regarding multiple meteorological variables. Subsequently, these data are used in models to create the weather forecasts. As such, models play an essential role in everyday life.

Models are simplified representations of reality (Frigg & Hartmann, 2024; Refsgaard, 1996; Savenije, 2009). They exist in different forms, such as mental, scale and computer models. Mental models are an idea, belief or concept within your own mind of how the world works. It is often formed by personal experience (Craik, 1967; Mayer et al., 2017; Moray, 1999). Many people use the mental model 'Margin of Safety' in their daily life – incorporating a buffer in your planning to be able to handle unforeseen circumstances (Thompson, 1979), for instance, adding more time for train journeys to take into account potential delays. Scale models are physical representations of an element of the world. In a scale model, the dimensions of reality are adhered to, but the size is often smaller (Black, 2019; Frigg & Hartmann, 2024). Computer models are codified representations of reality in quantitative terms. These are used to support calculations and predictions (Savenije, 2009; Devi et al., 2015). Within my thesis, I focus on this last group of models.

Computer models are used throughout science and practice, for many applications and within many disciplines. For example, climatology (Eyring et al., 2019; Knutti et al., 2013), ecology (Schmolke et al., 2010; Quinn et al., 2017), epidemiology (Amorim & Cai, 2015; Garner & Hamilton, 2011) and hydrology (Burt & McDonnell, 2015; Horton et al., 2022; Refsgaard, 1996) use models frequently. Applications include improving process understanding, exploring future scenarios, such as climate scenarios, or creating real-time simulations, such as weather forecasts. Burt & McDonnell (2015) found that, based on a bibliometric analysis, the use of models steadily increased over time, up to 70% in 2010. Besides models being applied for scientific purposes, models are also used for practical purposes. For instance, models are applied for real-time flood monitoring, or for predictions of droughts. In the practical context, models are used to inform decision making in water management.

1.2 | Water management and modelling

Water management's goal is to govern the water resources available sustainably and safely (Brouwer, 2015; Cosgrove & Loucks, 2015). In the past, and still now, this has not always been done: groundwater has been overabstracted, river systems have been disrupted and water bodies have been polluted (Connell & Grafton, 2011; Cosgrove & Loucks, 2015; Darling, 1945; Nabavi, 2018; Zwarteveen et al., 2021). More recently, water managers and scientists recognise that all water bodies require careful management (Bartholomeus et al., 2023; Brouwer, 2015; Connell & Grafton, 2011; Cosgrove & Loucks, 2015). How water management is specifically arranged differs per country. Generally, a national organisation keeps an overview of the national needs and demands, while local water management organisations deal with the local needs, plans and implementation of water management strategies. Within water management, different priorities can be considered, such as agriculture, environment and urban areas.

Hydrological problems are embedded in society (Blackett et al., 2024; Jakeman et al., 2006; Wescoat Jr & White, 2003), which means they have a mutual dependence. For instance, urban areas need to implement and maintain proper water systems to cope with heavy rainfall or waste water (Otterpohl et al., 1997; Xu et al., 2020). Or, hydrological extremes will become more frequent due to climate change (Brouwer, 2015; Connell & Grafton, 2011), which means agriculture will have to adapt to more floods or droughts (Anderson, 2014; Quinn et al., 2017). Also, water availability will be more uncertain, which can impact the environment (Hughes et al., 2011; Melsen et al., 2018a). These problems require decision making to handle them, which is often partially informed by model results – one way how modellers can potentially act as influencers, even extending to influencing society. However, these decisions are made with many uncertainties (Malano, 2010). Moreover, any decision regarding water management will have a socio-economic effect. For example, water rationing during a drought might devalue the environment or negatively impact agricultural enterprises (Anderson, 2014). Thus, decisions informed by model results can have far-reaching socio-economic impacts (Jakeman et al., 2006; Uusitalo et al., 2015).

As model results can have socio-economic impacts, model results can be questioned and criticised in this context, for example model results were questioned after flooding in the Brisbane area, Australia (Supreme Court of New South Wales, 2021; Malano, 2010). Therefore, model results need to be created responsibly, meaning that the model results can be explained and understood, for which its creators and/or users can be held accountable. Transparency and retraceability can increase the accountability of the model results. Transparency relates to being open about assumptions and decisions made throughout the modelling process (Klein et al., 2024; Malano, 2010; Packett et al., 2020). Retraceability means that how model results were created is identifiable in hindsight and by other actors (Lahtinen et al., 2017). However, transparency and retraceability are perceived differently per person, depending on individual experience and background. A modeller (in this thesis, defined as the subject-matter experts in representing the system being modelled using numerical relationships) might find a model to be more transparent than a policy maker, because they have a different background. Furthermore, model results, or any form of knowledge, can be perceived differently. Packett et al. (2020) explored how gender analysis can inform modelling decisions and influence model results interpretation. They showed that including other perspectives in a modelling study and the representations in a model redistributes the decision making to be more equitable and, as such, more responsible.

1.3 | Philosophies of science

Perceptions of knowledge – in this case, how model results are viewed and interpreted – are dependent on philosophies of science. Philosophy of science concerns itself with why and how science creates and delivers knowledge. It tackles the question of when something can be called scientific (Frigg & Hartmann, 2024). This type of philosophy has many different stances that each define differently what "good" science is. Examples of science philosophies are:

- **Positivism** — Logic and reason form the base for knowledge creation in this science philosophy. It poses that all knowledge based on observations and logic is objective. Within this science philosophy, there is an emphasis on quantitative methods (Guba & Lincoln, 1994).

- **Post-positivism** — Objectivity is also pursued in this science philosophy, however, it is also recognised that the scientist's background and values can impact what is observed. Post-positivists consider both quantitative and qualitative methods (Guba & Lincoln, 1994).

- **Constructivism** — This science philosophy is based on the idea that knowledge is actively constructed. It is created through the interaction between a scientist and all other aspects, such as the methods, society and the problem studied (Evanoff, 2004; Guba & Lincoln, 1994).

- **Subjectivism** — All knowledge is deemed to be subjective, meaning there exists no objective truth. Subjectivism poses that each scientist has their own understanding of knowledge and that knowledge is created from the scientist's individual perspective. Experience of the scientist is perceived to be important in determining reality (Smith, 1908).

- **Relativism** — Like subjectivism, relativism is the idea that no absolute truth exists. However, knowledge is created in relation to cultural, societal and historical influences. Relativism acknowledges and respects differences in culture, but does not provide a way to overcome them (Evanoff, 2004).

These five examples are part of a spectrum of possible philosophies: many more philosophies exist. Besides determining model result interpretation, philosophies of science colour the manner of communication about model results between modellers and between modellers and other stakeholders, such as decision makers. Additionally, scientists might not be aware of their own philosophy of science. And even if they might be aware of it, they can still shy away from it due to philosophy's ambiguity (Glattfelder, 2019; Laplane et al., 2019).

In the natural sciences, a positivist stance seems to dominate. In this context, objectivity entails that scientific results do not contain anything from the scientist themselves (Daston & Galison, 2007; Frigg & Hartmann, 2024; Leamer, 2009). Objectivity is shown through, for example, creation of images or figures, statistics, or double-blind clinical trials (Daston & Galison, 2007; Gustafsson, 1981). However, in the past few decades, questions arose concerning what 'scientific objectivity' actually is (Longino, 1990; Stefanidou & Skordoulis, 2014; Tsou et al., 2015; Wood, 2001). As such, objectivity is a changing concept through time.

The philosophy of science community became interested in modelling because models represent reality in an idealised manner, but are not completely true to it (e.g. Frigg & Hartmann, 2024; Hegselmann et al., 1996; Oreskes, 2003, 2018; Parker & Winsberg, 2018). As such, modelling undermines the traditional viewpoint that natural sciences give a 'true' representation of reality (Frigg & Hartmann, 2024). Even though modelling undermines the idea of objectivity in natural sciences, quantitative results often have the association of being objective. In modelling, scientists are moving towards acknowledging subjectivity and bias in its process more (Babel et al., 2019; Krueger et al., 2012; Melsen, 2022). To reiterate a previous example, Packett et al. (2020) showed how gender might influence the modelling process and model result interpretation, as such undermining the idea of objectivity. As such, philosophies of science can colour the way a scientist views and interprets the results of and uncertainties in their scientific process.

1.4 | Uncertainty in modelling

Throughout the process of modelling natural environmental systems, uncertainty is introduced. Three types of uncertainty can generally be discerned: epistemic, methodological and technical (Bastin et al., 2013; Nearing et al., 2016; Oreskes, 1998). However, Nearing et al. (2016) pose that uncertainty can only be understood based on a philosophy of science. For example, positivist thinking typically only acknowledges technological uncertainty. Technical uncertainty pertains to the traditional uncertainties studied: data, model structure and parameter uncertainty (Krueger & Alba, 2022; McMillan et al., 2018; Oreskes, 1998; Uusitalo et al., 2015). Not surprising, given the dominance of a positivist stance in hydrology, these sources of uncertainty are most frequently addressed in the hydrological modelling literature. Epistemic uncertainty pertains to a lack of knowledge or understanding regarding the characteristics of the process being modelled or the modelling process itself (Oreskes, 1998; Uusitalo et al., 2015). For example, lack of data or information can result in an incomplete understanding of the full system. Certain physical processes might not be apparent because of that, meaning that these processes might be excluded from the representation of the model. Methodological uncertainty refers to the method chosen and used (Bilcke et al., 2011; Mercieca & Mercieca, 2013). Applying a different method entails having a different representation of reality. It can be underdetermined which method would be most fit-for-purpose. Of these uncertainties, technical uncertainties can often be more easily quantified, which highlights the variation possible in model results.

A factor introducing variations in the model results is the variation in modelling practices, as choosing a method introduces differences in results (Clark et al., 2008; Melsen et al., 2019). Since the modellers choose the methods, (hydrological) modelling results depend on who executed the modelling study – another way how modellers can act as influencers, which I coined as inter-modeller variability. Holländer et al. (2009) invited ten scientific modellers to model the exact same German artificial catchment. None of the modellers had any prior knowledge of the catchment. All were given the same data. Despite this, the model results differed substantially, mainly due to the modeller's personal judgement and choices made during the modelling process. Holländer et al. (2014) extended this study by arranging a field visit and giving access to more data. The modellers were asked to redo the modelling study with this additional information. Again, Holländer et al. (2014) found that the modeller's personal experience played an important role in the differences in results. Therefore, to be able to put model results in perspective, it is important to understand how decision within the modelling process are made.

Modelling decisions are based on a multitude of motivations. Babel et al. (2019) explored how scientific modellers make decisions when developing an environmental model. They conducted seven in-depth interviews, and found that the team and the organisation in which the modeller works influences the decisions they make. Melsen (2022) conducted interviews with modellers in three different scientific hydrological modelling groups. She found that the team in which a modeller work is highly influential in making modelling decisions. The modelling decisions within the modelling process result in path dependency (Lahtinen et al., 2017). Decisions later on in the modelling process are limited by decisions made earlier. As a modeller makes these modelling decision based on their own context, such as experience, habit and work environment, inter-modeller variability occurs.

1.5 | Standardisation of the modelling process

Reverting back to water management, where models are used to support decision making. In Section 1.2, it was mentioned that models have to be developed responsibly. The inter-modeller variability introduced in the previous section can challenge this responsible development. Standardisation is one way to counter inter-modeller variability, and with that to contribute to accountability. Standardisation, in a general sense, aims to create a process or product of similar quality and with the same features in a similar way (Cambridge Dictionary, SA). Within modelling, this translates to unifying modelling processes and practices through, for instance, frameworks or guidelines. Standardisation can be implemented differently, e.g. voluntarily, through certification, or through legal requirements, and can be executed with various approaches, e.g. through automation, or guidelines (Deltares, 2023b; Goulden et al., 2019; Jones et al., 2020; Lord & Anthony, 2000).

Standardisation can have multiple advantages. It can create consistency, transparency, reproducibility and retraceability of modelling results. Consistency relates to executing a modelling study in a similar manner (Goulden et al., 2019; Müller et al., 2014). Standardising the modelling process results in a similar way of working between modellers. Transparency and retracability can be increased, because assumptions and decisions between modellers are the same with standardisation, meaning these only have to be documented once. Reproducibility is the ability to replicate modelling results without any unforeseen errors or differences (Gundersen, 2021; Schwarz et al., 2020). These advantages can contribute to a better support base for decision making based on models (Wang et al., 2023). However, whether these advantages come to fruition depends on how the standardisation approach is implemented.

Standardisation also has some potential disadvantages, such as lack of flexibility and innovation, and the costs it incurs. As methods and processes are unified, flexibility is decreased (Meijer et al., 2023; Wears, 2015). This potentially decreases the modellers ability to change the process according to the specific needs of a modelling study. Innovation can be hampered because the standardisation approach is no longer questioned (Pardo-del Val et al., 2014; Wears, 2015). Also, modellers will no longer learn from each others modelling practices as these have been standardised. The standards might not align with the experience of the modeller on which approach works best. Finally, it requires resources, effort and time to implement a standardisation approach (Daniels & Starr, 1997; Jones et al., 2020). These disadvantages can result in standardisation becoming detrimental to creating a support base for decision making.

In water management, several standards have been explored. In the Netherlands, for instance, this has been done through the Good Modelling Practice handbook (van Waveren et al., 1999), through automation (Deltares, 2023b) and through the development of national frameworks (Hoogewoud et al., 2013; Kukuric, 2011). In Australia, a national modelling framework is developed as a way to implement standards (eWater Source Group, SA). Each of these approaches bring their own challenges and opportunities. Since it is the modellers that have to implement the standardisation approaches in their everyday modelling practice, it is crucial to understand how modellers perceive different standardisation approaches.

1.6 | My positionality

In this thesis, I will research modelling practices and standards and how modellers perceive these – in essence, I reflect on how modellers reflect on their process. I will interpret these results, and since a scientist's perspective determines how data are interpreted and communicated, it makes sense that I, with a more post-positivism outlook, also reflect on my own perspective. Therefore, I include a positionality statement, which contains how my personal background has impacted my choices for and within this research.

I am a Dutch woman, 28 years old (at the time of writing this statement, Sept 2024). Already from a young age, I loved water: seeing it, playing in and with it (for any stories ask my mum). Being near the water clears my head and calms me down. Water and how it flows have always fascinated me So, for me, it made sense to start my studies in climate and hydrology and follow this up with my PhD research in the hydrological sciences.

Besides water, I like numbers, I like calculating things, preferably out of the top of my head or with pen and paper as my aids. As such, I find the idea of models interesting (even though they are on computers). I also like a certain amount of clarity and being able to explain things. These aspects draw me quite naturally towards the positivist philosophy, which my mainly natural-science education has strengthened.

Also, I have always loved stories, discovering a different world. This translated to curiosity in other cultures and perspectives. My interest in different perspectives has grown into a more post-positivist outlook with probably unconsciously a hint of subjectivism and relativism. My PhD allowed me to explore the perspectives of many people on modelling and also meet many more people to hear their stories.

I completed my BSc and MSc at Wageningen University, respectively in 2017 and 2020. During my studies, I focused on hydrology and climate, learning about both social and physical aspects of these domains. My dual interest in social and environmental sciences meant I found this PhD a great combination of the two. My research was within the hydrological disciplines and trying some modelling (which failed though), but applying interview techniques and bringing a social perspective to it.

This is the positionality I had at the start of my PhD. I will reflect on how this has changed during and due to my PhD research in **Chapter 6**.

1.7 | Thesis outline

Within this thesis, I focus on decision-support modelling in water management, analysing the modeller's practices and their perception of standardisation approaches. In a decision-support context, model results are used in situations with high societal impacts. Therefore, understanding the social aspects in the modelling process – how modellers can act as influencers within it – can contribute to putting modelling results into perspective, and help create a more accountable modelling process. The main research questions that will be addressed in this thesis are:

1. How do hydrodynamic decision-support modellers make modelling decisions?
2. How do hydrodynamic decision-support modellers perceive automation as a standardisation approach?
3. How do decision-support modellers perceive top-down and bottom-up standardisation approaches?

In the first three research chapters of this thesis, the research questions are explored in detail. In **Chapter 2**, it is discussed how modellers motivate their modelling decisions. These results are based on fourteen interviews conducted with Dutch hydrodynamic decision-support modellers. **Chapter 3** relates how modellers perceive automation as a standardisation approach. Automation as a standardisation approach means replacing manual modelling decisions with a computer programming script that executes those decisions without human intervention. Centralised automation has the potential to ensure quality of modelling decisions. **Chapter 4** contains the analysis on how modellers perceive two different standardisation approaches. I address two case studies: a top-down standardisation approach currently implemented in Australia and a bottom-up standardisation approach that exists in the Netherlands.

Chapter 5 summarises the findings and insights I obtained during this research into recommendations for the hydrological modelling network – all actors, i.e. funders, commissioner, modellers, users, decision-makers, involved in and influencing the modelling study. Finally, in **Chapter 6**, this thesis's findings are summarised and discussed in a broader context and asses how the modellers can act as influencers within the modelling process and in society. This also includes a reflection on my positionality with regards to this research topic.

This thesis highlights the social aspects prevalent in hydrodynamic decision-support modelling. Although the responsibility of and differences in modelling results are often ascribed to the individual modeller, this thesis shows that the context in which a modeller works is just as, if not more, important in influencing the model results.

2

A modeller's fingerprint

This chapter is based on:

Remmers, J.O.E., Teuling, A.J., Melsen, L.A. (2024). A modeller's fingerprint on hydrodynamic decision support modelling. *Environmental Modelling & Software*. *181*, DOI: 10.1016/j.envsoft.2024.106167

Abstract

MODEL results can have far-reaching societal implications, requiring fit-for-purpose models. However, model output is resulting from a particular path chosen with each modelling decision. We interviewed fourteen modellers in the Dutch water management sector in order to study how decision-support hydrodynamic modellers make modelling decisions. An inductive-content analysis was performed. We identified eight motivation-categories. Individual and team considerations mostly motivate modelling decisions. We identified patterns between the motivation-categories and their occurrence across modelling steps. Modelling decisions during model implementation were found to be more in the modeller's direct sphere of influence, while decisions concerning model structure and data selection more outside of it. So, even though modellers can leave their fingerprint, their sphere of influence and thus their fingerprint's clarity is bounded by institutionalised predefined decisions. Thus, models and their results are shaped within a broader sphere than the modeller's alone, requiring a broader consideration of organisations and standards.

"One man's way may be as good as another's,
but we all like our own best."

—Jane Austen, *Persuasion* (1818)

2.1 | Introduction

After flooding in January 2011 that afflicted large areas near the Australian cities of Brisbane and Ipswich, a lawsuit was filed against two dam operating companies and the state of Queensland. The claim was that negligent operation of the dams, and not the rainfall directly, had resulted in widespread property damage. In the lawsuit, the decisions concerning early dam releases, the modelling they were based on, and the handbook that was followed were questioned. It was concluded that the flooding occurred due to a combination of torrential rains in the catchment and a release from the Wivenhoe Dam to prevent it from overtopping. The judge, initially, ruled in favour of the claimants, but the ruling was later overturned during an appeal (Supreme Court of New South Wales, 2021). As this case shows, model results and decisions informed by model results can can have real consequences and the role of the modeller can be scrutinised.

Model results used for decision support can have far-reaching effects and model users, the decision makers, can be held accountable, as the Australian case shows. This means that model users depend on modellers to provide them with accountable model results. However, model results can vary depending on the path taken with every modelling decision that is made (Glynn et al., 2017; Holländer et al., 2009, 2014; Lahtinen et al., 2017; Melsen et al., 2019; Polhill & Edmonds, 2007). We refer to *modelling decisions* as the choices made throughout the modelling process by the modeller, including: how and which processes are represented, how data are used and how the validation is executed. Each decision has a reasoning behind it, which is what we call the motivation for this modelling decision. Given that the modeller is in charge of many of these decisions, it is imperative to take the modeller's role in the modelling process into account to understand their role in the modelling process.

Modelling decisions made by a modeller can impact the various stages in the modelling process. As detailed in **Chapter 1**, Holländer et al. (2009, 2014) showed that the modeller's personal judgement contributed substantially to the variations in the model results. Also, Melsen et al. (2019) conducted a modelling experiment to test how four modelling decisions impact the modelling results. In stead of having other modellers make modelling decisions and execute the modelling, they defined the modelling decisions and their options when they conducted the modelling study themselves. The studied modelling decisions significantly affect the simulated flood and drought events. These studies show that model results depend on the modelling decisions made. With these modelling decisions, modellers leave their fingerprint in their models.

Several studies already investigated how social processes steer motivations for modelling decisions, mainly in an academic setting. Melsen (2022) conducted interviews with hydrological modellers to investigate how they made their modelling decisions during a specific modelling study. Her conclusion was that the team in which a modeller works is an important factor in making certain decisions in a particular way. Babel et al. (2019) also conducted interviews, yet across different disciplines and focusing more on the development of the model. Also here, it was found that the modeller's team and collaborators contributed considerably to model development. Addor & Melsen (2019) looked at how model structures are selected. They performed a bibliometric study, in which they found that legacy, represented through the institute of the first author, is the best predictor for model selection. These studies show a wider range of motivations in modelling decisions, beyond the considerations of the individual modellers, and thus the large role of social processes in technical model use.

However, decision-support modelling generally has its roots outside of academia, at governmental agencies and consulting companies. There are several studies who have investigated the modeller's influence in such a decision-support setting. Padilla et al. (2018) present a survey regarding modellers' perspectives on modelling and simulations. Their survey was fully completed by 151 respondents who were identified as model builders from both academia and industry. They found that conceptualisation and validation of models relied heavily on informal methods, such as the use of pen and paper or visual inspection. This implies that individual modeller's perspectives shape the modelling decisions made. Deitrick et al. (2021) deployed a survey, which had 27 respondents, and conducted four in-depth follow-up interviews, which focused on how modellers make decisions during the watershed modelling process. They reached an audience working in mainly academics and governmental agencies. Their study highlights how the modeller's ethical – related to personal standards – and epistemic – related to knowledge building – values inform the modelling process. Fleming (2009) conducted a small survey about how a watershed model is selected. Their respondents worked in government, the private sector and academia. They found that both individual considerations, such as familiarity with a certain method, and organisational considerations, such as standards in an organisation, are considerable influences in the modelling process. These studies show that the social aspects of modelling are important to consider in different contexts – academia, government and the private sector.

Here, we explicitly and solely focus on practitioners that use models for every day decision-making practice. Understanding modelling decisions in such a context is extremely relevant, because these models directly interfere with the real world (Lane, 2014), as also demonstrated by the role of models in the Australian flood. We conducted fourteen in-depth interviews with modellers at water authorities and consulting companies, focusing solely on modelling in the governmental and consulting sectors. The analysis consists of an inductive content analysis covering the motivations behind modelling decisions, the variation of motivations across the modelling process and the difference in motivations between the governmental and consulting sectors.

2.2 | Methodology

In order to investigate how decision-support hydrodynamic modellers make modelling decisions, we conducted fourteen interviews. We used the Netherlands as a case study, which will be described in the first subsection. The interviewees worked at local water authorities or consulting companies. The interviewee selection and interviews are detailed in the second subsection. The analysis of the interviews consisted of an inductive content analysis, which is elaborated in the third subsection.

2.2.1 | Case Study

Since social processes are highly localised, it is infeasible to further our understanding of the social aspects of modelling at a general level. Therefore, we focus on a defined case: decision-support modelling in the Netherlands. Such an approach is also defended by Deitrick et al. (2021), having conducted a survey and interviews in the Chesapeake Bay Watershed (US), indicating that other case studies would improve the general understanding of how decision-support modellers make modelling decisions.

In the Netherlands, national and local water authorities and consulting companies are the main parties that execute decision-support modelling for water management, assisted by research institutes. At the national level, the governmental agency Rijkswaterstaat has this responsibility (Government of the Netherlands, SA). At the local level, there are 21 water authorities, which carry responsibility for the water management in their region (Government of the Netherlands, SA). Over the past decades, the management structure of Rijkswaterstaat and the water authorities has shifted. Initially, in the 1950s and 60s, they had substantial in-house modelling knowledge, which allowed them to facilitate their own modelling studies. At some point, the structure changed from having in-house knowledge to performing project management. As a result, Rijkswaterstaat and the water authorities had to rely more on consulting companies and research institutes to (partially) carry out the modelling process (Vukovic, 2022; van den Berg & van Lieshout, 2022). Through public biddings, there are several consulting companies that support the water authorities. This support consists of executing of the whole modelling process, setting up the model, or knowledge provision and training for water authorities so that they are able to set-up and execute models themselves. Research institutes support the hydrodynamic modelling mainly through developing new model software. The research institutes specifically target the knowledge intensive and technical aspects of the software development. Moreover, the research institutes play a key role in the maintenance and quality assurance of the software suites. Some of the interviewees of this study that work at water authorities indicated that, since a few years, water authorities aim to obtain more in-house knowledge again through human resources and knowledge acquisition.

For decades, most water authorities have worked with the same model software suite: SOBEK (Deltares, 2023d; Stelling & Duinmeijer, 2003). The functionalities of this software suite cover among others rainfall-runoff processes, 1D open and closed hydrodynamics and 2D overland flow. Currently, the water authorities are looking into suitable alternatives to SOBEK, as its maintenance is discontinued.

2.2.2 | Interviewee selection and interviews

To study the modeller's motivations for their modelling decisions, we interviewed fourteen modellers at water authorities and consulting companies. The interviewee selection was within a hydrodynamic modelling project and through snowball sampling, which is not an unusual method. Still, it could affect the representativeness of the interviewee sample. We examined this by evaluating the saturation of coding and the differences between the interviewees working at the same organisation.

Nine of the interviewees worked at six different water authorities at the time of the interviews. The other five worked at four different consulting companies. Modellers at water authorities and consulting companies execute other aspects of the modelling process due to the different role of each organisation. The interviewees were all hydrodynamic modellers. Their experience with modelling ranged from one to fifteen years. The interviewees use models for various applications. For flood applications, all interviewees execute real-time forecasting to evaluate if their water system can cope with certain events. Other applications include infrastructure dimension design and scenario testing, which were both mentioned by about half of the interviewees. Additional, though less-frequently mentioned applications are water system design, drought management and water quality modelling. Figure 2.1 provides an example of modelling decision the interviewees encountered during their modelling studies.

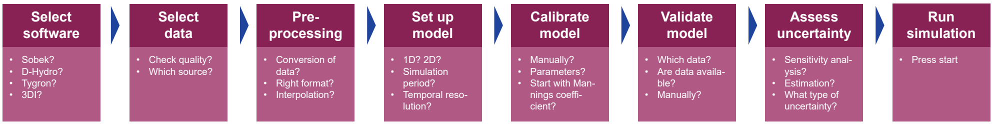

Figure 2.1 | General overview of the modelling process and examples of decisions that modellers have to make for each step, for Dutch water management modellers.

All interviews, semi-structured, took place between September and December 2021. All but one were conducted in Dutch, with the other one being held in English. On average, the interviews lasted between 1 hour 15 minutes and 2 hours. At the start of the interview, all interviewees gave their informed consent. After the interviews, we sent the transcripts for a final approval to the interviewees, which was granted by all of them. The interview guide, included in Appendix General, covered questions about motivations behind modelling decisions, i.e. how the interviewee made a certain modelling decision or when they would change their decision. For instance, one of the interviewees considered the aim of the modelling study to determine the simulation period: for one flood event a couple of days or weeks and for groundwater-related studies at least a year. Each interview was recorded and subsequently transcribed.

2.2.3 | Content Analysis

An inductive content analysis of all transcripts was carried out using AtlasTi, version 9 (ATLAS.ti Scientific Software Development GmbH, 2022). This analysis, first, entailed inserting topical interview codes to categorise which segments of the transcripts covered the different modelling decisions and which segments covered general questions, such as the interviewee's confidence in a model and its simulations and the interviewees' perceived influence as a modeller. For example, the topical interview code 'Calibration' indicates all quotations in the transcript that were about calibration. This can be in response to a question about calibration, but it can also be that the interviewee mentioned calibration in relation to another question.

The second part of the content analysis produced interview codes, which contain the results of this study – how modellers make modelling decisions. Here, we specifically focused on whether and how modellers make a certain decision and when they would decide differently. The set obtained from this inductive content analysis contained 96 interview codes (see Appendix of Chapter 2). Each interview code can be applied to multiple quotations – an excerpt of the transcript. A single quotation can also have multiple codes. For example, we assigned two codes ('Tender requirements' and

'Earlier work within organisation') to the following quotation: 'The decision was made based on a proposal that was tendered and awarded before I started at the company'. The code 'Tender requirements' relates to the fact that it was based on a tendered proposal and the code 'Earlier work within organisation' covers that this proposal was tendered before the interviewee started working at the organisation.

After the inductive content analysis, the interview codes were classified into motivation-categories, which resulted in a general overview of the motivations. This classification was made through discussion within our team. In our analysis, we have also looked at overarching themes. These themes appear in several of the previously determined categories. Furthermore, combining the categorised codes with the topical codes (i.e., topics) allowed for exploration of different motivations for different modelling steps. For example, the motivations in the category Organisational appear on average most frequently in the quotations related to the topical code 'Model structure', which implies that decisions related to model structure generally occur at the organisational level. Moreover, we analysed the average occurrence of codes, compared between modellers at water authorities and consulting companies. Since water authorities and consulting companies have a different role in the Dutch water management, i.e. water authorities are responsible for decision making and often using models to support that, while the consulting companies support the water authorities in executing the modelling process, we hypothesise that this leads to different motivations in making modelling decisions.

2.3 | Results and Interpretation

2.3.1 | Motivation classification

Across all interviews, 1699 quotations that contained a motivation were identified in total. These quotations could be classified into 96 different motivations (Appendix of Chapter 2), which were again grouped in eight different motivation-categories (Figure 2.2). A modelling decision is the choice for a certain method within one step of the modelling process. The reasoning to choose this method is the motivation for this decision. A decision can have multiple motivations, also from different categories. The eight identified categories are: Individual, Team, Organisational, External, Commissioner, National, International and Consequential. These are discussed below. Although we identified these eight categories, some interview codes could not be classified in one category, therefore we classified them in two. For example, 'Testing' is in both the Individual and Team category, because this is sometimes executed by the individual modeller and other times it is executed within the team.

Individual — This category refers to all motivations made at the discretion of the modeller themselves. This category is the biggest category of all eight in terms of both codes and quotations, even when not taking into account the shared codes with the Team category. The most prevalent codes in this category are 'Personal insight' and 'Personal experience' with 152 and 115 codes, respectively. An example of how personal experience influences a modelling decision is the following: *"I know from experience that they [statistics-based cross sections] are not always reliable*". This interviewee chooses to not use certain data because in their experience the approach of statistics-based cross sections is not reliable enough. Another frequently-occurring interview code is 'Personal preference' with 62 quotations. Other codes included in this category are among others 'I don't know', 'It doesn't make a difference' and 'Logic'.

Team — The motivations in this category relate to the modeller and their direct colleagues. For example, a decisions is made by discussing it within the team or with their superior. The most-frequently occurring code in this category is 'Experience colleagues', of which an example is *"I think it is more because my colleagues also do it in that way, so that is kind of easy. But I get what you mean, you could also retrieve the roughness values in a different way. For me, it is actually because it was done in this way here, so I have adopted that method*". This modeller adopted a certain method, because their colleagues used it and had experience with it. There is considerate overlap with the Individual category, in total 261 distinct quotations, divided over twelve interview codes. Among others, the interview codes 'Testing', 'Model run time' and 'Hydrological knowledge' are the main codes that overlap, respectively accounting for 77, 59 and 42 interview quotations. For example, 'Testing' can be executed by an individual modeller or a team.

Organisational — The organisation a modeller works at can also influence the decision they make, for instance through the vision the organisation has or through the infrastructure it provides. Another motivation in this category is 'Earlier work done in organisation', of which an example is given in this quote: *"Well, we provide a certain tool to other organisations. (...) We also use this tool a lot.*". Here, an organisation developed a certain tool for other organisations. Because of this work done, they also use it frequently themselves. This category shares two codes with other categories. One of them is 'Best available' with the Team category. This refers to what is perceived to be the best available method or data within either a team or an organisation.

External — Within the modelling process, tasks can be outsourced to an external partner. If this external partner is responsible for making the modelling decision, the motivation has been classified in the External category. This might mean that the interviewed modeller is unaware of the full reasoning for certain decisions made by the external partner, which is generally the case when an external partner executes these modelling decisions without in-between consultation. We captured this in the code 'Executed by an external partner'. This motivation was assigned to 37 distinct quotations. This happens regularly for data pre-processing, which is often already performed by the research institute providing the data. For instance, a modeller uses meteorological data for flooding assessments. These generally-available statistical data or time series have been derived by the Royal Netherlands Meteorological Institute, and it is in this instance not necessary for the modeller to do any additional pre-processing: *"No, we don't do that [pre-processing of meteorological statistical data or time series]. No, that is actually already validated by KNMI [the Royal Netherlands Meteorological Institute].*". Another task that is commonly executed by an external partner is the 'Model set-up', calibration and validation for the main model of a region. A consulting company supports a water authority in this way. Still, the water authority can and will adapt this main model when deemed necessary. Of course, the execution of any modelling step can also be an iterative process between two organisations, including discussions between the parties. This is reflected in the code 'In discussion with external partner', occurring 23 times.

Commissioner — This category includes the motivations that are influenced by the commissioner of the modeller. The modeller would often be the external colleague or partner of the commissioner. In this sense, External and Commissioner are the other sides of the same coin. In total, this category comprises seven codes, including for instance 'Commissioner determines' (37 quotations) and 'Requirements for model study' (13 quotations). One of the important motivations in this category is 'Time available in project', which the following quote highlights: *"Yes, that [calibration requirements] is very much in consultation, so it depends a bit on the effort you've put in, so the total amount of hours versus the timeframe that is there. Often we say: 'Guys, we have put in this much effort, everyone has looked at it, the money has run out' and at that moment it [the calibration] is simply cut off*". In this quote, they choose to finish the calibration and continue with the modelling process and accept that this will be the best achievable calibration in light of financial and time resources. This category share one code with the Organisational category: 'Limits costs'.

National — Modelling decisions can also be bounded by national considerations. This could be national laws that modellers have to adhere to, or the origin of a tool or method that modellers use. The origin of a tool might ease the use of it, because the documentation or support might be provided in the modeller's first language. Additionally, another motivation example is 'Generally used', which means that a particular method is used across multiple or all organisations in the Netherlands: "*I think also because it is the most used model software within the Dutch hydrological community*". This might have grown organically, as it is easier to compare model studies when using similar methods, or the various organisations came to an agreement on what to use.

International — This category indicates motivations that are based on international factors, such as international agreements. This category is the smallest of all, as most modellers were focused on regional hydrological support modelling. We identified two types of international aspects that modellers used in their modelling decisions: first, data sharing, and second, agreements regarding rules for environmental protection. The first aspect entails water authorities using data from neighbouring countries (i.e. Belgium or Germany) to obtain more data that is potentially of a higher quality. This most often applied to forcing data. The second aspect relates to constrictions imposed by especially the European Union. As an interviewee has highlighted with regards to European environmental goals that need to be taken into account in the design of the water system: "*And also the goals you want to achieve. So, there are all kinds of environmental objectives that are imposed from Europe. And how should I say it, extreme is not the right word... The more ambitious they are, that of course also influences your scenario. So when they say 'Oh no you can just have some boring grass', then you don't have to heavily wet that area or anything. But when they say 'There must be a swamp here and this and that', then of course you have to make all kinds of adjustments to raise those groundwater levels. Those kind of things*". This in turn impacts what needs to be modelled.

Consequential — Consequential refers to a choice being (partially) predetermined because of an earlier choice made, demonstrating path dependency. For example, within particular modelling software only certain model settings are available. Then the choice for that software package limits the choices for the settings. For example, an interviewee mentioned "*I think the default value.*" regarding the choice of the simulation time step. In this case, the interviewee needed to select a maximum temporal resolution that the model cannot exceed. For this maximum, the interviewee used the default value provided by the model. Another aspect important in this category is that parts of the modelling process are executed automatically: the modeller chooses a tool to automatically perform certain tasks in the modelling procedure, but then the choices in the automation tool are a consequence of the choice to use that tool.

Some of the interview codes could not be classified in any of these eight categories. These interview codes pertain to the availability of data or software or technologies, attributes of certain methods or data or model software, and specifics of the model study itself. Because the interviews were set up in a generic manner, the interviewees

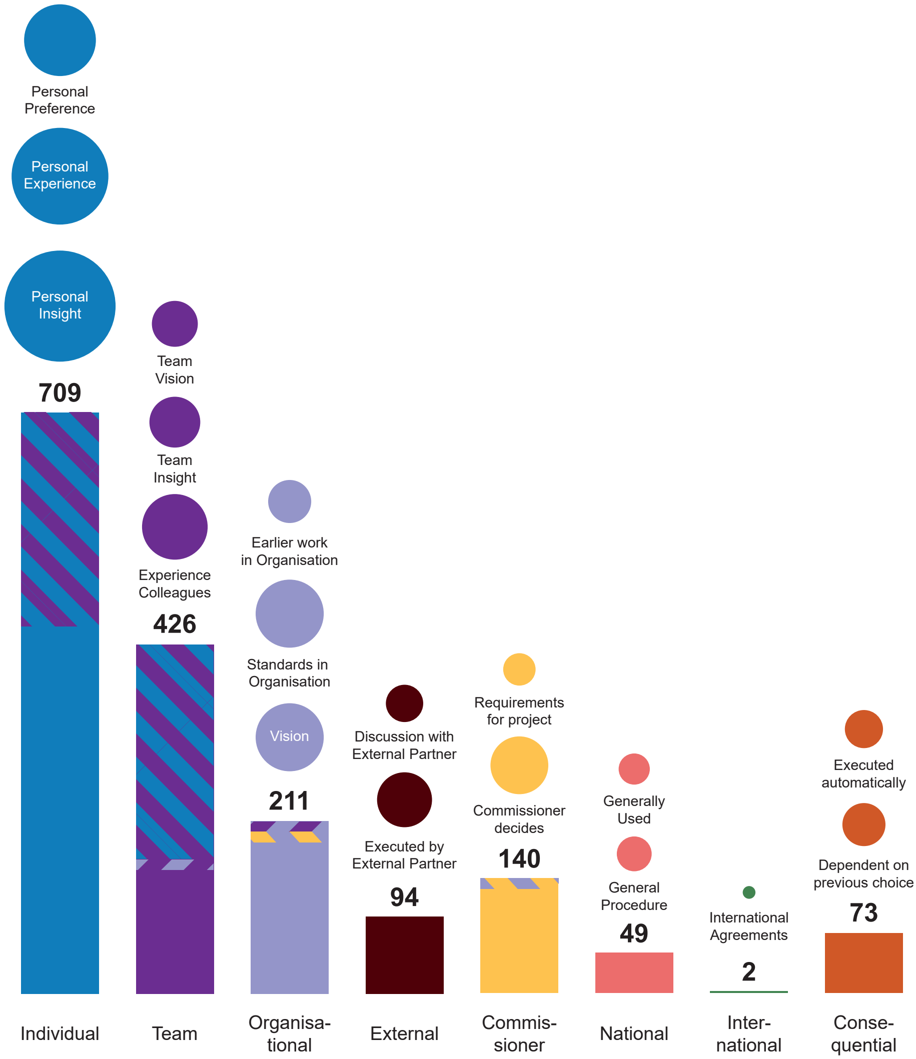

Figure 2.2 | Classification of motivations behind modelling decisions in hydrodynamic modelling for water management. For each motivation-category, one or more examples of specific codes are given. The number of quotations of the example codes is indicated with circles, which are area-proportional.

specified often that their decisions depended on the goal of the model study, the model structure used, the project, the study area and the circumstances. The interview codes 'Based on theory', 'Maintenance stops' and 'Model stability' are stand-alone codes.

2.3.2 | Overarching themes

Aside from the eight categories, we have identified two overarching themes within the interview codes: Vision and Standards. These themes are recurring with distinct interview codes across several categories. An example are the codes 'Team vision' and 'Vision organisation', which are respectively classified in categories Team and Organisational. Both codes relate to vision. Some codes that were not classified in any of the eight previously introduced categories do relate to either of the themes and are discussed in this section.

Vision — Vision, as defined by the Cambridge Dictionary, means 'the ability to imagine how a country, society, industry, etc. could develop in the future and to plan for this' (Cambridge Dictionary, 2024b). In our study, we look at vision related to how the modelling process is shaped and how modelling decisions are made. This vision is informed by the values an individual modeller, team or organisation holds, and can also be created in each of these categories. A certain vision can lead to a preference for a method, however, preferences are not necessarily visionary. An overview of all codes related to vision and the frequency with which they occur is visualised in Figure 2.3. Codes related to vision occur in three different categories: Individual, Team and Organisational. In the Individual category, the vision relates to a single modeller's perception on how the modelling process should be executed. In our study, this is only visible in three quotations within the code 'Personal preference'. Only these three quotations have been visualised in Figure 2.3. In a team, vision is visible by the team vision they set out. For example, an interviewee said that the team they work in created a certain vision on which modelling software they wanted to use. They created this vision because it would be easier to work together and fill in for each other. Within this vision, the modellers, including the interviewees, work as much as possible with the model software they determined.

At the organisational level, the organisation creates and prescribes their vision. This vision can relate to which model software is used, how the model is set up (e.g. 1D or 2D schematisation), or what data is used. For instance, one of the interviewees indicated how the vision of a water authority determines which model software is used. The interviewee follows this vision too, among others because it works well. This vision was initially formulated by the hydrological modellers and approved by

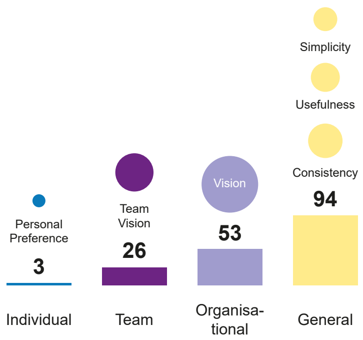

Figure 2.3 | Different categories, and examples of codes per category, that all relate to the overarching theme Vision. This theme indicates that certain modelling decisions are made because they fit a certain modelling vision. The circles, which are area-proportional, represent the number of quotations for a code. The quotations in the Individual category are a subset of the code 'Personal preference'.

the board. Afterwards, everyone at the water authority carries out this vision.

On top of that, some codes related to vision could not fit in one of the previously defined categories and are labelled as general vision. These codes include among others 'Simplicity', 'Consistency', 'Usefulness' and 'Efficiency'. Simplicity, for instance, can mean that a modeller wants to keep the modelling process 'as simple as possible, but as complex as necessary'. One of the interviewees described that they as an organisation require consistency in the results, even when different parties have generated the results. Some of these general aspects of vision can be ascribed to one of the eight different categories dependent on the context. For example, a modelling team can strive for simplicity in its modelling process, because that aligns with their vision of the modelling process – fitting the Team category. Also, a modeller can value efficiency and tailor their decisions to that – fitting the Individual category.

Standards — Standards are either a generally accepted method or a level of quality (Cambridge Dictionary, 2024a). In our study, we refer to standards as: a standard is a generally-accepted (informal) or prescribed (formal) method or way of working. The interviews showed that standards are implemented in three different categories: Organisational, National and International. Most standards that were mentioned were formally documented. An organisation can have internal standards or generally used procedures, exemplified in this quote: *"There are standard models from which I cut out a part. Those models have been created once, I do not know how long ago*". This interviewee uses (parts of) the general model to execute their specific modelling study, which is an informal standard. This quote also highlights a disadvantage of a generally-accepted approach or standard procedures: evaluation and updates might not happen regularly once accepted and trusted.

In the National category, there are several ways in which standards are implemented, for example based on a national handbook (the Good Modelling Practice guidelines, see also **Chapter 4**) or on national laws. An interviewee gave another example: the different organisations agreed on using particular data. They recorded this agreement in a guideline. Organisations have to adhere to some international laws as well, for example as described for the International motivation category, which have to be taken into account in how the modelling process is executed.

The codes in these three categories – Organisational, National and International – are the visible aspects of standards, that is the modeller is executing (part of) the modelling process is aware of them. There are probably also standards in the categories External and Commissioner. However, this is not visible to the modeller, since the modeller is an outside party in both cases. Standards could also exist at the team level, but were not encountered as such in the interviews. Possibly because these standards have grown more organically or are easier internalised. The use of standards will be more elaborately explored in **Chapter 3 and 4**.

2.3.3 | Distribution of categories across different modelling steps

While being cautious in quantifying qualitative data, we identified some patterns between the eight identified motivation-categories and how often they occurred across the different steps of the modelling process (Figure 2.4). We analysed how one category is divided over all modelling steps (Figure 2.4a) and how all categories are represented within each modelling step (Figure 2.4b). We have left out the category International from this part of the analysis, because it only comprised two quotations.

Figure 2.4a shows that, in general, the modeller makes decisions based on their individual and team motivations during each step of the modelling process. However, the personal and team motivations are more strongly featured in the modelling steps related to model implementation – i.e. from modelling step 'Model set-up' to 'Validation'. These are the modelling decisions in the modeller's direct sphere of influence – the aspects a modeller can directly change. The categories Organisational, External, Commissioner and National, outside the modeller's direct sphere of influence, feature more frequently for motivations behind decisions about 'Model software and Data selection'. These modelling steps are often formalised in the organisation's vision or available infrastructure. The category Consequential is used as motivation across all modelling steps. Since most modelling steps are related to each other, it is unsurprising that this category is a motivation in all modelling steps. However, this category occurs considerably more in the modelling steps 'Pre-processing' and 'Model set-up'. For the 'Model set-up', this is because this modelling step is partially dependent on the default settings in the model software. This thus implies that some decisions in the 'Model set-up' are not explicitly made, but a consequence of other decisions made earlier.

Figure 2.4b depicts the division of the categories within each modelling step, which can indicate what type of motivation tends to inform a decision in each modelling step. Figure 2.4b shows a similar general pattern as Figure 2.4a – Individual and Team motivations inform the modelling steps concerning model implementation, while the other motivation categories inform the 'Model software and Data selection' more. Still, Figure 2.4b highlights some other details too. For instance, the category External accounts for about a quarter of the motivations in the modelling step 'Pre-processing'. This is in line with how forcing data are retrieved by the modellers: they retrieve them already pre-processed from the Royal Netherlands Meteorological Institute. This institute is considered an external partner that executed the pre-processing. It is also shown that the external partner and commissioner influence the modelling decisions 'Sensitivity analysis', 'Calibration', 'Uncertainty analysis' and 'Validation'. This can be explained by the general outsourcing of these steps from water authorities to the consulting companies due to a higher computational capacity and experience with these modelling steps available at the latter. In summary, both figures clearly depict that certain modelling decisions are more likely to be in the modeller's sphere of influence than others. The modeller makes decisions in model implementation, but is often confined in using a particular model or particular data by the organisation or at the national level.

2.3.4 | Differences in motivations between water authorities and consulting companies

Another way to divide the motivations is based on where the interviewee worked, a water authority or a consulting company. The difference in responses between these two organisations is compared for each category we identified. We perceived no considerable difference for six of the categories: Individual, Team, External, National, International and Consequential. For the two overarching themes, there was only a substantial difference for the theme Standards. In Figure 2.5, the average number of quotations per interview are shown for the categories Organisational, External and Commissioner and the theme Standards.

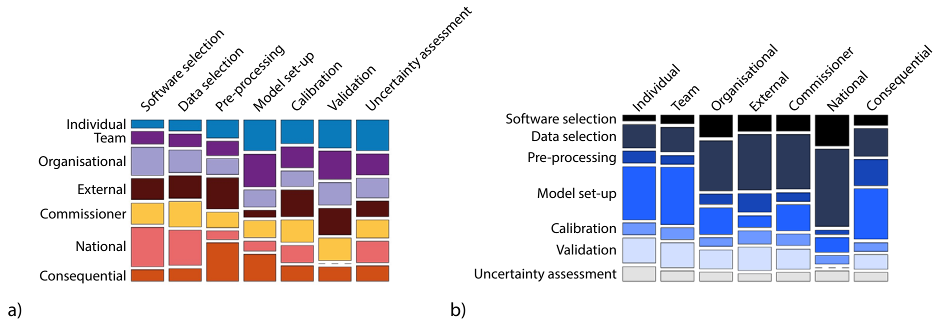

Figure 2.4 | Distribution of motivation-categories across different modelling decisions. a) How the motivation-categories are divided within each modelling step. b) How one motivation-category is spread across the modelling steps. The motivation-category International was excluded, because it only contains two quotations.

The interaction between water authorities and consulting companies can explain the differences in the average quotations per interview. Water authorities are often the commissioner of modelling studies, while consulting companies generally execute (parts of) of the modelling process for a client, in this case the water authorities. In line with this, the interviewees at the consulting companies more often mentioned motivations within the Commissioner category. An interviewee at a consulting company mentioned that the commissioner (i.e. a water authority) would provide the data in order to ensure consistency between modelling studies within the water authority. Interviewees from water authorities mention some codes classified in the Commissioner category, because these relate to project management (e.g. 'Limit costs' and 'Time available in the project'). Where the consulting companies adhere to the vision and requirements of their commissioner, the water authorities define their own vision and requirements within their organisation. Hence, the water authorities have a higher average number of quotations in the category Organisational. Still, modellers at consulting companies are influenced by the organisation they work in, for example through the organisational infrastructure or internal standards. For the difference in the category Standards, this has to do with the visibility of the standards. Modellers at water authorities see the standards their organisation imposes. Modellers at consulting companies generally see the standards that a water authority asks for as, for example, 'Determined by commissioner'. We saw this in the individual codes of the Standards category: the interviewees at water authorities mentioned 'Standard in organisation' substantially more than interviewees at consulting companies. The other codes were more evenly distributed, because modellers at consulting companies also adhere to the national standards or have a similar perception that an option is not a

Figure 2.5 | Comparison of motivations between water authorities and consulting companies.

choice (anymore).

We expected that motivations from the External category would be considerably more frequent for the water authorities, since they outsource some of their modelling. As said before, the External category represents in some way the other side of Commissioner. However, there is no substantial difference between the water authorities and the consulting companies in this category. We do see that within this category there is a difference that aligns with our expectation: the code 'Executed by external partner' occurs more frequently for the water authorities than the consulting companies. Water authorities use consulting companies (their external partner) to execute parts of the modelling process in the Dutch water governance system. Consulting companies also use external partners, mainly in data pre-processing. This is why they scored equally with water authorities in this category.

2.4 | Discussion

2.4.1 | Modeller's Sphere of Influence

With this study we investigated how hydrodynamic modellers in Dutch water management make modelling decisions. Other studies have conducted similar work regarding the social aspects surrounding modelling, but focused on different aspects. Our results indicate that most modelling decisions are made at the individual level (the category Individual had the highest number of interview quotes). This contrasts the findings of Melsen (2022), who found that most modelling-decisions are made at the team level. We hypothesise that this can be explained by how the interviewees conduct their modelling work. Whereas the interviewees from Melsen (2022) worked mainly in large research teams and concerned themselves more with model development and scientific publications, our interviewees worked in smaller modelling teams and were more involved in model application. This difference in context can explain why the individual category is larger in our study compared to the study of Melsen (2022) and can also explain why different classifications emerged from the data compared to Melsen (2022). Babel et al. (2019) – an interview study about how models are constructed in various disciplines – also recognised that the team and collaborators are key actors in model development, which was seen across the different disciplines. These studies (Babel et al., 2019; Melsen, 2022) highlight team considerations as a main influence on modelling decisions and within model development.

Other studies, that also include perspectives from government and industry, recognised that individual, team and organisational considerations play a considerable role in modelling decisions. Fleming (2009) found that non-technical issues – issues related to the context in which a modelling study is executed – account for 27 % of the reasonings in model selection. Within this 27 %, the organisation, costs and standards within the industry were the top three motivations. This aligns with our results, in which the national level and the organisation were influential in model selection. Deitrick et al. (2021) show that ethical values, alongside epistemic ones, are used in making modelling decisions. Generally, more epistemic values were mentioned. However, Deitrick et al. (2021) recognise that their respondents related 'that values are not something that they typically reflected upon'. One of our interviewees brought this up as well: at the end of the interview they expressed that the interview had been insightful for them too, since they do not have a lot of time for reflection during their day-to-day tasks. Another interviewee reflected that it used to be possible to trace back a certain model to an organisation or sometimes even a particular modeller based on its model structure and settings. This clearly reflects that modellers leave a fingerprint on their models. Other studies have also shown the modeller's fingerprint on the model results (e.g. Holländer et al., 2009, 2014; Krueger et al., 2012; Lahtinen et al., 2017; Saltelli et al., 2020). All studies show that modelling decisions are within the modeller's sphere of influence, meaning that a modeller leaves their fingerprint on the results.

2.4.2 | Institutionalisation and Internalisation

Even though a modeller leaves a fingerprint on the model result, other factors impact the modelling decisions too. Melsen (2022) introduces the concepts of institutionalisation and internalisation. Institutionalisation occurs when a team or organisation takes up an individual modeller's method as a general method. Babel et al. (2019) recognise that 'methods can be actants in shaping organisations', which reinforces the concept of institutionalisation. This shaping is due to the development of certain infrastructure. Internalisation means that an individual modeller makes the methods used in their team or organisation their own. Babel et al. (2019) use the concepts incorporation and anchoring. Incorporation means that a choice of a certain method is transferred from one person to another, including the process of making this method your own. Addor & Melsen (2019) have alluded institutionalisation and internalisation, as well. They highlight that continuous use of the same model creates a particular 'modelling ecosystem', i.e. institutionalisation. Organisations might be more prescriptive about model structure and data selection to ensure a baseline for modelling quality. The modelling outcomes have to be accountable and reliable when used in decision support.

Internationalisation was not explicitly observed in the interviews of this study but was hidden in the answers of the interviewees. For example, as described for the Team motivation category, one of the interviewees stated that they adopted the common method in their organisation with regards to the retrieval of roughness values. This interviewee was aware of other alternatives, but found it easiest to adopt this particular method. An interviewee highlighted internalisation by knowledge being passed down from one modeller to the next: *"On the one hand, that [sensitivity analysis] is passed down so to say, from hydrologist to hydrologist. And what works is just shared"*. Babel et al. (2019) refers to this as embedded social knowledge – knowledge is passed down and adopted.

Institutionalisation was more visible in our interviews. The creation of standards within an organisation is a form of institutionalisation. An organisation can formulate a certain workflow or prescribe the use of certain data. Moreover, the code 'No longer a choice' encompasses the result of institutionalisation: a modeller no longer feels as if they have a choice because a standard is present in the organisation. An interviewee indicated that at one point a choice was made and this is now copied by them. This standard in the institute can be perceived as if there is no longer a choice. Similar to our code 'No longer a choice', Babel et al. (2019) highlighted that for the modellers they interviewed some decision seemed 'evident', either because of popularity, standards, typical or common use. Internalisation, institutionalisation and modelling decisions seeming evident imply that motivations are generally applicable to various modellers or within a discipline.

However, generally-applicable motivations are in contrast with modellers leaving a fingerprint on the modelling results. Babel et al. (2019) also saw a pattern that interviewees distanced themselves from generalisations. Initially, their interviewees started with the generalisation 'Everybody', which was slowly lessened to their own discipline and to subgroups within their discipline. We distinguished a similar pattern, where our interviewees preluded their answers with 'It depends on ...'. The aspects it depended on were goal, study area, model structure, situation and project. Other times, our interviewees referred to a distinct example of when they made a certain modelling decision. This break-away from generalisations supports the idea of a modeller leaving an individual fingerprint. Even though a modeller can leave a fingerprint, their sphere of influence and thus the clarity of their fingerprint is bounded by institutionalised predefined decisions.

2.4.3 | Fit-for-purposeness

The tendency to link motivations to specific circumstances implicitly means that according to the interviewees a model should be fit for purpose. A fit-for-purpose framework was developed by Hamilton et al. (2022). They define three requirements for a modelling study to be fit-for-purpose: Usefulness, Reliability and Feasibility. Each requirement covers a different context, respectively end-user and management, problem and project context. Hamilton et al. (2022) have indicated multiple key considerations per requirement. Taking these into account, the motivations from our interviewees do seem to align with the requirements (as indicated in Appendix of Chapter 2). For example, Usefulness covers among others codes related to 'Depends on ...' and the category Commissioner. Reliability is seen in the codes 'Hydrological knowledge / processes', 'Logic' and 'Testing'. The codes 'Limit costs', 'Available time in project' and 'Personal experience' are represented by the requirement Feasability. This suggests that modelling in the Dutch water governance system seems to align with this fit-for-purpose framework. However, the codes from our interviews mainly fall in the Feasibility requirement (866 quotations, compared to 340 for Usefulness and 283 for Reliability), while in the fit-for-purpose framework it is recommended that the motivations are more balanced across the three requirements (Hamilton et al., 2022).

In the context of model usage for decision support, which require fit-for-purposeness, the interaction between the modeller and the decision maker is relevant. Just as modellers have their own values about, perspectives on and expectations of models, so do decision makers and other stakeholders (Borowski & Hare, 2007; Deitrick et al., 2021; Hamilton et al., 2019; van Voorn et al., 2016). We did not cover this aspect in our interviews, however, our interviewees did address this sometimes. Some interviewees mentioned that they did not perform an uncertainty analysis, partly because they experienced previously that decision makers did not know how to handle uncertainty ranges in their decision making. Yet, addressing model uncertainty is part of the reliability requirement of the fit-for-purpose framework. Therefore, creating fit-for-purpose models should be a joint effort of all stakeholders, in which they will carve their modelling path together.

The interaction between stakeholders is necessary before, during and after the modelling process. However, certain intentions might not be realised. The realisation of a modelling study can be described as a path, on which multiple decisions are made at forks. The interaction between stakeholders can ease the retracing of the modelling steps if necessary (Lahtinen et al., 2017). The retracing is based on checkpoints, peer review and other forms of evaluation. Lahtinen et al. (2017) provide specific recommendations. Still, as one of our interviewees indicated, evaluations are currently not formally executed on a regular basis, especially during crisis situations. They mentioned that during the modelling study not enough time and funding is available to execute evaluations. Also, the time for evaluation is after a crisis, but even then it is not often executed due to other pressing matters. In the Australian example, also a crisis situation, they did adapt the operators manual, so future situations would be handled differently (Supreme Court of New South Wales, 2021). This does show hindsight evaluation. With evaluation in place, a perfect modelling path is not guaranteed, however, following a poor path can be avoided (Lahtinen et al., 2017). Our interviewees seemed to be willing to have a more adaptive modelling approach. To put these evaluations in place requires commitment from modellers to be as transparent as possible, from decision makers to have the conversation about uncertainty, from the commissioner to provide the infrastructure and from stakeholders to be willing to engage throughout the modelling study.

2.5 | Conclusion

In this study, we explored motivations behind modelling decisions for hydrodynamic decision-support modellers in the Netherlands. We conducted fourteen interviews with modellers from water authorities and consulting companies. Afterwards, we executed an inductive content analysis on the transcripts. The analysis lead to a classification of modelling decision motivations with eight categories: motivations based on individual considerations, team considerations, the organisational level, external inputs, the commissioner's requirements, the national level, the international level and consequential effects. Additionally, two overarching themes were identified: Vision and Standards. Furthermore, we evaluated which category of motivations dominated for different modelling steps. On top of that, we looked at differences in modelling motivations between modellers from water authorities and from consulting companies.

Our results indicate that most modelling decisions are made at the individual level (the category Individual had the highest number of interview quotes). Mainly decisions related to model implementation are within the modeller's sphere of influence – the aspects an individual can (in)directly change. This is where the modeller can leave a fingerprint: one interviewee indicated they were able to recognise which modeller created a certain model schematisation. Most of the model software and data selection is based on motivations in the categories Organisational, External, Commissioner and National. These aspects tend to be outside the modeller's direct sphere of influence. Still, modellers do see that modelling decisions depend on the context of the modelling study, implying that a model should be fit-for-purpose. The motivations in our case study seem to align with the requirements (Usefulness, Reliability and Feasibility) of fit-for-purpose, but in our case, feasibility seemed to be more of an argument than reliability and usefulness. This means that other factors, such as institutionalised predefined decisions, limit the modeller's sphere of influence and thus the sharpness of their fingerprint

3

A modeller's perspective on automation

This chapter is based on:

Remmers, J.O.E., Teuling, A.J., Dahm, R.J., van Dam, A., Melsen, L.A. (2024). Power to the programmer: Modeller's perspective on automating the setup of hydrodynamic models for Dutch water authorities. *Socio-Environmental Systems Modelling*. *6*, DOI: 10.18174/sesmo.18657

Abstract

USE of models in decision making, for example in water management, requires confidence in the model and its outputs. Since choices in model setup affect model output, this confidence is affected by the modellers' professional judgement. Computer programmers can use their expertise in coding to standardise some of the tasks associated with computational modelling. Therefore, centralised automation has the potential to ensure quality of modelling decisions. Since it is the modeller that makes the choices in the model set-up, it is important to understand how modellers perceive automation. To explore their perspectives, we conducted fourteen interviews with modellers at water authorities and consulting companies in the Netherlands. The transcripts were analysed through deductive and inductive content analysis. Our study reveals that automated modelling processes can improve efficiency, transparency and consistency, but only if certain requirements are met, such as good documentation, clear ownership, adequate maintenance and frequent evaluation. Therefore, managing the risks and benefits of automation requires balancing the power between modellers and programmers.

"But it is his way.
One man's style must not be the rule of another's."

—Jane Austen, *Emma* (1816)

3.1 | Introduction

As already elaborately discussed in **Chapter 1 and 2**, models are increasingly used in many diverse disciplines (Eyring et al., 2019; Garner & Hamilton, 2011; Schmolke et al., 2010) and for a wide range of applications. Besides their academic use, model results also support decision making. The decisions informed by models can have major effects on people, increasing or decreasing the risk of damages or loss of life. Therefore, it is important that people have confidence in the model and the outputs that serve as the basis for decision making.

Modellers influence the modelling process through the decisions they make in setting up the model. Their influence includes questions of value, matters of fact, ease of use and uncertainty (Babel et al., 2019; Lim et al., 2023; Melsen, 2022; Voinov et al., 2018). Some examples of modelling decisions, as defined in **Chapter 2**, throughout the modelling process are given in Figure 3.1. Examples of specific decisions made are the parameter value of the levee height or model boundary and resolution choices, such as the grid size. With each modelling decision made, a path is created behind each model result (Glynn et al., 2017; Lahtinen et al., 2017; Melsen, 2022; Moallemi et al., 2020; Polhill & Edmonds, 2007), meaning that it is crucial to consider how and in which context the model is used.

Technologies, including models, contribute to shaping the context and modelling decisions made (Verbeek, 2008; Latour, 1990; Melsen et al., 2018b). Take this example: a choice in a model was hardcoded by a colleague years ago. Another modeller uses this model later and subsequently relies on the method implemented by their colleague. This influences their future options within the modelling process, for example, by having constraints on the spatial resolution they can choose. Conceptual representations and particularly software suites are created for specific purposes, meaning that care must be exercised when models are use in new domains or for new purposes. (Hamilton et al., 2022; Beven, 2000). As such, which model is chosen and how it is used by the modeller is important.

The responsibility that comes with the use of models consists of setting up models based on several key principles: appropriateness and transferability, reproducibility and transparency (Crout et al., 2008; Zurell et al., 2020). Appropriateness and transferability refer to the potential of a model to function well in novel situations. This is directly linked to the purpose of the model: often models have better transferability only within a specific purpose for which they were created. Therefore, it is important to delineate and communicate the model's purpose (Werkowska et al., 2017). Reproducibility refers to the capacity for another person, team or organisation to redo the modelling done by someone else (Gundersen, 2021; Schwarz et al., 2020) without leading to unforeseen or unsolvable errors or differences, and ideally, resulting in identical model outputs. Transparency refers to the interpretability and understandability of the modelling process, including but not limited to communication of assumptions, uncertainties and motivations underlying modelling decisions. Transparent modelling reveals the path behind the modelling results, which involves values and biases. The implementation of these principles in modelling for decision making contribute to justifying model results and their trustworthiness (Wang et al., 2023). For the practical implementation of these principles, different approaches can be used.

Standardisation, as defined in **Chapter 1**, is one approach to justify the model setup and use (Jakeman et al., 2006). Standardisation can streamline the modelling process, increasing efficiency, reproducibility and transparency (Howard & Björk, 2008; Müller et al., 2014; Schmolke et al., 2010). It can also create consistency between modelling studies (Müller et al., 2014; Vrontis, 2003). This facilitates the possibilities for model evaluation through intercomparison. However, standardisation also decreases flexibility, for example, by preventing a model's capacity to capture regional differences. By definition standardisation depends on the acceptance of a single perspective on how a system should be represented, which could lead to path dependence (Wears, 2015). Standardisation can also result in reduced innovation (Mir & Casadesús, 2011). How the standardisation is implemented affects how and which advantages and disadvantages will surface. Furthermore, the standardisation's implementation and the model purpose influence whether the advantages outweigh the disadvantages of standardisation.

The implementation of standardisation can vary widely, ranging from voluntary methods to more prescriptive methods. An example of a voluntary method is the creation and use of guidelines, such as the 'Good Modelling Practice' handbook in the Netherlands (van Waveren et al., 1999, , see also **Chapter 4**), the Australian Groundwater Modelling Guidelines (Barnett et al., 2012) or the Traffic Modelling Guidelines for the UK (Beeston et al., 2021). More restrictive methods include incorporating modelling requirements into law (Fan et al., 2018) or issuing certifications for adhering to certain practices (Balci, 2001; Sampaio et al., 2011). Any of these standardisation methods might include the automation of the modelling procedure.

Automation is a broad term that can be defined in multiple ways. Because we are particularly interested in the execution of computational models, in this study, we define *automation of the modelling process* as the replacement of manual modelling decisions with a computer programming script that executes those decisions without human intervention. Automation in this sense is frequently used in the pre-processing of data and in calibration procedures. If programmed well, automation decreases the chance of human error. Automation also increases modellers' efficiency, by reducing their cognitive load (Lewis et al., 2018).

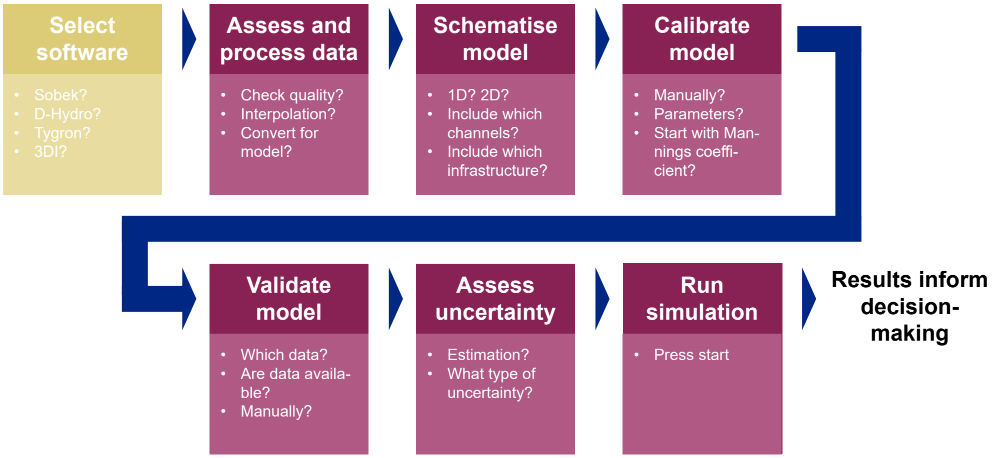

Figure 3.1 | Simplified representation of the modelling process. Each rectangle is a modelling step, in which multiple modelling decisions can be made. Below each step, we mention some examples of modelling decisions that the modellers that we interviewed encounter in their modelling work. Automation of the modelling process refers in our study to all step except 'Select software', because this step was often not in the sphere of influence of the interviewed modellers. After the simulation is run, the results are used to inform decision making by the decision maker, which is often another person than the modellers themselves. This last aspect is not part of this study. This representation of modelling steps is based on Figure 1 in Refsgaard et al. (2007).

Automation introduces a paradox. Initially, because automation requires explicit representation through computer scripting, transparency of the modelling process may be increased—the steps being automated need to be made explicit, and rationales for the automation need to be made clear to programmers. Yet, over time, automation can also obfuscate the underlying processes if certain requirements, such as documentation of the automation procedure, are not met. Also, automation removes decision power from the modellers, shifting this power to the person or group that develops the automation script – i.e. 'the programmer' (note that one person can embody both roles at different moments in time). Even though automation is always a form of standardisation, we refer to standardisation as a way for decision makers to control the quality of modelling studies and their underlying modelling decisions. Since modellers are subject matter experts that make the modelling decisions and shape the model results, it is important to understand how they perceive the trade-offs between potential efficiency and transparency gains and shifts in power that may be associated with the act of automation. Figure 3.1 depicts stages of the modelling process that might include automation that we will examine in this paper (all steps except 'Select software').

In this study, we explore modellers' perceptions on automation with a case study of decision-support hydrodynamic modelling in the Netherlands. We conducted fourteen interviews with modellers working at Dutch water authorities or consulting companies that participated in a project that included automation of modelling decisions.

3.2 | Methods

3.2.1 | Case Study

Section 2.2.1 provides a detailed description of the case study used also in this chapter. Figure 3.2 demonstrates a common distribution of tasks between water authorities and consulting companies based on the interviews for this study. In this study, we focus on automation developed and used by local water authorities and consulting companies.

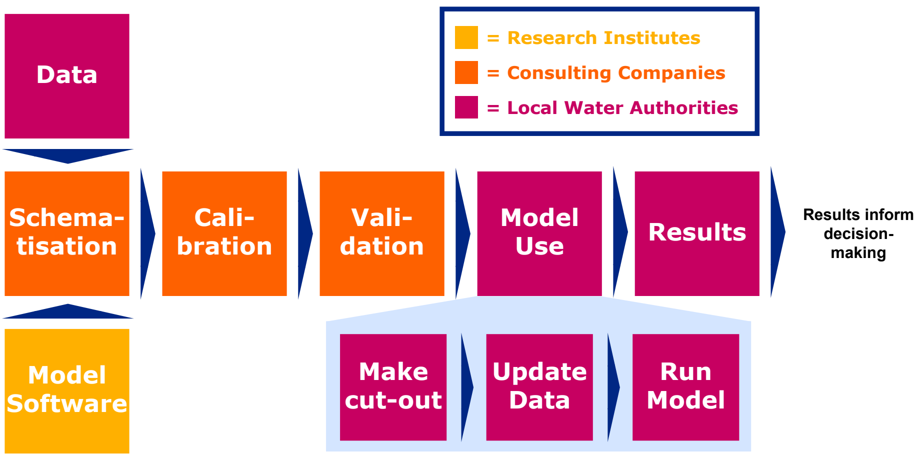

Figure 3.2 | Representation of a general modelling process for decision support in managing the Dutch river systems. It depicts the frequently observed roles in Dutch water management. Local water authorities, consulting companies and research institutes execute different modelling steps. In this study, we focus on the local water authorities and consulting companies, because our interviewees, at the time, worked at either one of these organisations. The process might look different depending on context and vision of the organisations involved. Furthermore, the process often involves an iterative procedure (not indicated for clarity).

Over the last decades, most water authorities used the same model software for rainfall-runoff processes, 1D open and closed hydrodynamics and 2D overland flow: SOBEK (Deltares, 2023d; Stelling & Duinmeijer, 2003). Automation scripts to set up and use SOBEK already existed locally and were partly incorporated in the model software. A new modelling suite, D-HYDRO Suite 1D2D (Deltares, 2023a), is developed, eventually leading to the discontinuation of updates of SOBEK. This instigated the transition to D-HYDRO for many of the water authorities. An important development in tandem is the HyDAMO dataframework, a shared data structure for the water authorities (Nederlands Hydrologisch Instrumentarium, SA). The switch to D-HYDRO is accompanied with an effort to centralise and unify already existing automation scripts scattered around different companies and organisations to facilitate the modelling process and stimulate the transition to D-HYDRO.

Centralising and extending the existing automation scripts was the goal of the HYDROLIB project, led by Deltares (Deltares, 2023b). This project was funded by a public-private partnership, which grants funding to consortia focusing on obtaining knowledge and innovation (a so-called TKI-project). Within this project, automation scripts from participating organisations are adapted for general usage and new scripts are developed and stored in a central public library, see Deltares (2023c). These automation scripts tackle different aspects of pre- and post-processing in the modelling process. The project also consisted of testing the scripts, mainly done by water authorities, and organising workshops for modellers to familiarise themselves with the automation scripts. These workshops were co-organised by the organisations that developed the automation scripts (the consulting companies and Deltares). The project ran for two years and finished in 2022.

The project results were picked up further in 2023, when the organisation behind the Netherlands Hydrological Instrument – a collection of data sets, model codes, tools and models to describe and simulate the full hydrological cycle in the Netherlands – organised HYDROLIB community meetings and joint script-developing working days. Additional functionality was added by several of the partners. The Netherlands Hydrological Instrument organisation has also commissioned 'review' training, where colleagues from different organisations will gain the knowledge to review additions or changes to automation scripts. Model code contributions require review before being added to the HYDROLIB code repository with the aim to assure quality standards. Hydrodynamic model generation for international applications is being implemented in 2023, building upon the general framework of HydroMT for automated and reproducible model building and analysis (Eilander et al., 2023).

3.2.2 | Interviewee selection and interviews

To explore the Dutch modeller's perspective on automation, we conducted interviews with modellers at the water authorities and consulting companies that participated in the HYDROLIB project. The interviewees, the same as in **Chapter 2**, were selected through convenience and snowball sampling. Most of the interviewees participated in the HYDROLIB project and might, therefore, be at the forefront of taking up new developments in modelling - or at least they are open to that. Therefore, we expect that they have a more positive outlook on automation than other modellers might have, which might bias our sample.

The fourteen semi-structured interviews were conducted between September and December 2021. The interviewees gave their informed consent at the beginning of the interviews. The interview guide is included in Appendix General. For the analysis in this chapter, we used a different part of the interview transcripts, also used in **Chapter 2**.

3.2.3 | Theoretical background

Part of our content analysis was based on the issues and best practices of automation derived from literature by Pagano et al. (2016). They described three issues and seven best practices for automation. They obtained their findings from other disciplines, such as meteorology, but apply these to automation in hydrology.

The three issues identified by Pagano et al. (2016) are: changes in the modelling process, in people's behaviour, and in the perception of the model's trustworthiness. First, modelling processes change due to automation, which may require the modeller to assume a different role or execute tasks not part of their expertise. Commonly, the first aspects that are automated are the most-easily automated tasks, such as data preparation. At some point, this leaves tasks that may not suit the modeller's expertise, which were previously done by someone else or not at all. Second, automation changes people's behaviour. Automation transfers the source of human error from the modeller to the programmer of the automation script. For example, inconsistent assumptions can be recorded in an automation script by the programmer. This transfer of errors can influence what a modeller communicates. For example, when the automatic system does not indicate that a warning should be issued, but the modeller estimates differently, what should the modeller communicate? If the modeller does issue a warning, they might have less ground to stand on depending on the operational protocol. Third, the perception of the trustworthiness of a model might not be accurate, e.g. due to a persistent first impression that is no longer valid. A modeller can perpetuate this first impression by recommending or dissuading the use of certain automation scripts. The issues with automation that were identified by Pagano et al. (2016) present themselves at different moments in the automation process. Since the case study we explore started recently, some issues cannot yet be observed.

Besides the three issues described by Pagano et al. (2016), they also derived seven best practices from meteorology. We focus on five of the seven best practices, because these centre on the link between modeller and automation. The other two best practices, 'Use automation to quality-control and ingest data' and 'Use well-designed forecasting interfaces', cover which aspects are most suitable to automate and how the automation should be designed, not the interaction between the modeller and the automation procedure. The five best practices we explore in this study that Pagano et al. (2016) suggest for the modeller are:

- **Have transparent systems** — A modeller should be able to retrace the different steps done by the automation procedure, such as intermediate results.
- **No peeking at the answer** — A modeller should give an estimation of the expected outcome before analysing the automated output. This creates better understanding of the results and makes the modeller more critical of the results from the automation.
- **Evaluate your results** — Evaluation of the automated output is crucial for improving the automation procedure itself. This should be done in a constructive and structured manner.
- **Never stop learning the science** — A modeller must keep learning the underlying science behind the modelling. This expertise will help the modeller in evaluating the automated results. This cannot be learned from modelling itself.
- **Redefine your role** — The role of the modeller might have to be redefined. As the modelling process will become less time consuming for the modeller due to automation, this might give the modeller more time for other tasks. However, this requires a different expertise, which might lead to the modeller being more removed from the modelling itself. This can reduce the modeller's ability to interpret and understand the results.

3.2.4 | Content Analysis

The transcripts were analysed in ATLAS.ti, version 9 (ATLAS.ti Scientific Software Development GmbH, 2022), through deductive and inductive content analysis.

Deductive content analysis — The deductive content analysis is based on the issues and best practices described by Pagano et al. (2016) (Appendix of Chapter 3A). Each issue and best practice was subdivided into several interview codes, resulting in 18 predetermined interview codes. This was directly based on the text in Pagano et al. (2016). For example, the best practice 'Evaluate your results' was subdivided in the interview codes: 'Evaluation', 'Evaluation decrease' and 'Evaluation increase'. 'Evaluation' covers anything that is related to evaluation of the scripts. The other two relate to how the evaluation develops over time. 'Evaluation increase' meaning more in-depth evaluation over time and 'Evaluation decrease' less.

Inductive content analysis — To supplement the predetermined codes of the deductive analysis an additional set of 89 interview codes (see Appendix of Chapter 3B) were developed through grounded theory in a first round of coding. These interview codes were divided over seven code groups, indicated in this appendix (column 'Group'). We were not able to group all codes, so seven codes were added to the 'Miscellaneous' group.

3.3 | Results and Discussion

First, we describe the findings related to the different levels of automation development and uniform use (results of inductive analysis). Second, we relate the results to the theoretical framework of Pagano et al. (2016) (results of the deductive content analysis). Finally, we discuss the different roles of water authorities and consulting companies and how this influences automation perception (results of inductive content analysis).

3.3.1 | Social levels of automation practice as standardisation procedure

The HYDROLIB project is about gathering and centralising automation scripts. The uniform application of these scripts contributes to increasing standardisation at the (inter-)organisational automation level. As such, this standardisation not only serves the purpose of increased efficiency for the modeller, but also as a way of quality assurance for decision makers (provided that certain implementation requirements are met - as will be discussed in this section).

The social levels of automation practice, presented in this section, are not necessarily directly related to the quality of the automation, but the scale is relevant to evaluate the usefulness of automation as a standardisation procedure. We will refer to these social levels of automation practice as automation levels. From the interviews it became clear that automation scripts already exist at several 'automation levels'. Based on that, we identified four levels at which automation is uniformly applied: Individual, Team, Organisational and Inter-organisational.

Individual — Any modeller will write an automation script at a given moment to create a more efficient and consistent workflow for themselves. This is often done for repetitive tasks that are cumbersome to execute manually and easy to automate. The scripts are built quickly, often having little to no documentation. For example, one respondent described their process as existing "*only on ad hoc basis at the moment that someone needs data for a part of the area. That would be executed with Excel-files that only the person who worked with it would understand*". The modeller is responsible for the development and implementation of their own scripts. Different modellers can create similar scripts, resulting in overlapping tools and inefficiency from the organisation's perspective. Another interviewee said it is sometimes more efficient to create your own script than to ask around if someone already has a script and figuring out how that script works. Here, the interviewee showed how a modeller considers the effort both options would take to make the choice to develop their own script. Of course, this would only be possible if a modeller has the capacity and knowledge to develop a new script. Evaluation happens ad hoc. At this level, the generalisable lesson learned is that automation as standardisation creates more efficiency for the individual modeller, but not necessarily more transparency and reproducibility between modellers in the modelling process, nor quality assurance for the decision makers. There is, however, a high level of agency for the individual modeller, flexibility to adapt the procedure and different approaches to the same task can provide insights in uncertainty introduced by choosing a particular method.

Team — Within a team, automation scripts can be created for common use. At this level, each individual maintains a degree of agency – the capacity to make their own decisions and be responsible for them. The scripts are developed and evaluated organically. Whenever a bug or a potential improvement is encountered, the script is adapted. The script often originates at the individual level, but has been proven to be useful for more than just one individual. Scripts are often used multiple times or as starting point for a specific application. Similarly to the individual level, evaluation happens impromptu and documentation is added only if necessary, but asking colleagues for help or explanation is easy. So in general, on the team level, standardisation in the form of automation results in more efficiency and collaboration between colleagues. However, this standardisation does not automatically result in more transparency as colleagues can easily ask each other for needed explanations instead of writing documentation.

Organisational — At the organisational level, scripts are required to be more generalised for internal use. Documentation is often added. In one example, an organisation established a monthly hackathon, in which an individual or team script was generalised for everyone to use. Such a hackaton or some other (formalised) approach or initiative is necessary to move a script from the Individual to the Organisational level. The organisation benefits because it can streamline its overall modelling and automation efforts. However, organisations can also encounter variation in approaches and the challenge of inter-operability between departments. This raises the question which method should be used and if an organisation-wide script is feasible and desirable. This is dependent on how this process is set up and how everyone is included in it. One approach raised by interviewees was that organisations established new processes for script development, for example the hackaton in the interviewee's organisations. This also raises issues about how evaluation of scripts would be handled. The evaluation of the automation scripts was dependent on how the scripts were developed and the vision and resources of the programmers. If scripts are developed for a commercial purpose, i.e. delivering a paid service to others, the organisation will have a financial incentive to maintain them, at least as long as the service is required. If the scripts are open-source, there is still the incentive to provide a good service, but the evaluation might also be partly the user's responsibility subject to the programmer's vision. For instance, Mer et al. (2020) recognise that in open-source modelling the documentation might become more fragmented due to the user's own responsibility. As such, it will partially depend on how willing and able individual modellers are in contributing to the evaluation and further development of a script. In summary, efficiency does increase in general with automation at the organisational level and there is potential for more transparency and reproducibility on this automation level. If automation is implemented properly with attention to maintenance, evaluation and documentation, this potential can be realised and contribute to standardisation and to quality assurance for decision makers.

Inter-organisational — Sometimes, different organisations each bring their own modelling culture, methods and scripts to coordinate inter-organisational automation efforts, such as HYDROLIB. In this circumstance, each organisation has to be open to a certain flexibility in their way of working, relinquishing some of their agency. Organisations involved in an inter-organisational initiative have to relay the outcomes of collaboration to their own modellers and programmers. These new or adapted tools might be in contradiction with what modellers are familiar with, meaning that the modellers might need time and incentive to adopt it. Bringing together these different scripts presents distinct challenges. For example, one interviewee stated, "*[when] steer-*

ing towards automation, ... the Waterschaphuis [Dutch governmental institution assisting all water authorities] [is] run[ing] into these problems, [for example,] within a water authority, [there] are problems of data formatting not align[ning] with what the hydrologist wants". The organiser of this effort also has to ensure that the automation is interoperable. This pertains to among others operating systems, programming languages, data sharing and computer capacity to run the scripts. The generalisable lesson learned is that automation at the inter-organisational level does motivate to unify data, enhancing the ease of data implementation and sharing. As such, automation as means of standardisation at this level can contribute to quality assurance for the decision maker through improved transparency, reproducibility and concistency. However, this does require the automation script to have, among others, documentation and an evaluation plan.

Figure 3.3 summarises the effects of standardisation across the different social levels of automation that emerged from our interviews.

Figure 3.3 | Depiction of change in the effects of standardisation when scripts move between levels at which automation is uniformly applied. Standardisation in this instance refers to ensuring quality assurance. These changes in effects of standardisation are from the (inter-)organisational perspective. On the side the requirements for standardisation are shown in order to achieve the effects.

As automation scripts are developed further, progressing from Individual to Inter-organisational level, the modelling process is slowly standardised to ensure quality for decision makers due to more reproducible and transparent modelling. This does necessitate among others documentation, evaluation, determined ownership and maintenance of the automation scripts. If these requirements are not met, the automation, as a standardisation method for quality assurance, loses its value. For example, centralised scripts from the (Inter-)organisational level can also be further developed at the Individual or Team level without feeding this back to the (Inter-)organisational level. This leads to decreased standardisation from an (inter-)organisational perspective. Additionally, automation at the (Inter-)organisational level may decrease insight into uncertainty, because it suppresses different methodological choices, and the freedom of the modeller to make choices. Therefore, modellers should maintain a good understanding of the hydrological system and the automation choices for the interpretation and evaluation of the results (Pagano et al., 2016; Woods & Sarter, 2000).

3.3.2 | Results based on Pagano's theoretical framework

Based on the interview codes informed by Pagano et al. (2016), two common topics were identified when analysing the interviews:

What is automated? — At the time of the interviews, mainly less complex components of the modelling process are automated within the Dutch hydrological community, at the individual, team and sometimes organisational level. The execution of modelling decisions (i.e. the choices made throughout the modelling process by the modeller) is automated, not yet the actual choices themselves. For example, the data importation into the model is automated, but which data and how is still determined by the modeller themselves. As such, the transfer of modelling decisions to the automation script does not occur frequently, according to the interviewed modellers. Still, the modellers are aware of giving power to the programmer when modelling decisions are automated beyond the individual level for purposes of standardisation. The modeller loses insight into the modelling decisions made by the programmer, which would also increase the difficulty to question and examine the assumptions behind the modelling results. Because of the loss of insight, modellers are hesitant to adopt automation at the (Inter-)organisational level, because they have to relinquish their own modelling decision power. For example, one interviewee stated, "*I, myself, would like to keep influence on that. But there will undoubtedly also be many people that just say 'Yes, it can be automated'. I want to know to some extent what the consequences are of the modelling decisions.*". Programmers, on the other hand, will receive more agency. The trade-off in agency affects how automation is perceived and illustrates the potential reduction in the modeller's responsibility as a consequence of increased automation (Limerick et al., 2014).

Convenience and efficiency are important factors to consider in Organisational or Inter-organisational level automation, even though this might lead to less transparency. For example, an interviewee stated, "*For ease of use, I would say yes, but I do think that you lose insight into certain [modelling] decisions.*". Decreased transparency can also result in a lower degree of interoperability of the automation scripts, especially if the effort is not made to understand the metadata. Multiple interviewees indicated that they would not read texts of tens or hundreds of pages long.

Besides a general loss of insight into the process, the loss of insight might vary per modeller, because each modeller has their own expertise with which they can understand and use the automation scripts. The sense of loss of understanding may also depend on the purpose of the modelling study, the resources available and the experience of the modeller themselves.

Appreciated practices by modellers — The interviewees mentioned several practices that they would appreciate and which would enhance the use and acceptance of centralised automation. First, transparency of the automation system is key, just as Pagano et al. (2016) mention in their best practices. This would also increase the ability to gain insight into the assumptions behind the automation. Transparency could be facilitated in many ways. One possibility mentioned by an interviewee would be to create an automatic log file of all the modelling decisions made with the automation. However, the log file is created at the end, which might not be used or read by the modeller. This renders the log file to be a silent form of communication between automation and modeller. Another option provided by an interviewee was creating pop-ups of characteristics of the automation when a modeller is using the automation. For instance, a pop-up might provide the limitations of a part of the model software if it is used or not. Carver & Turoff (2007) argue that an automation system should not be silent. This is also mentioned by Woods & Sarter (2000), who stress that people and technology are dependent on each other. This reinforces the idea of creating pop-ups or an equivalent. Yet, pop-ups would disrupt the modeller's work they do simultaneously to the automation, which affects their behaviour, for instance experiencing stress or impacting their performance (McFarlane & Latorella, 2002). Though McFarlane & Latorella (2002) do recognise human's ability to handle these interruptions, this depends on the extent to which the interruptions divert, distract, disturb and disrupt the current task. Besides, with pop-ups, modellers might just click '*OK*' without actually reading or considering the message. Another suggestion was that a slow build-up of the automation system would improve the automation's acceptance. First one aspect is automated, once modellers are familiar with this, a next aspect is automated. The slow build-up ensures that modellers can familiarise themselves with the whole system, leading to better understanding and acceptance of the automation. For the modellers, this will enhance their understanding of the system, which Pagano et al. (2016) recommends. The slow build-up also applies to getting their superiors on board, for example, an interviewee stated, "*How much time will you invest [in the script] the first time? And how certain is it that we will use it again? So, even though as a modeller you think you truly need this, often you have to convince your director that it is truly necessary despite the initial high costs.*".

3

Translating this to the inter-organisational level requires clear communication to the potential users of the automation script regarding the features of the automation, also recognised by Calder et al. (2018). During and after the script development, this means having good documentation, even though it might be time-consuming (Polhill & Edmonds, 2007) and the documentation has to be fit for purpose (Müller et al., 2014). However, it is not taken into account whether the documentation is read and used or not. Aside from the documentation, a central point for questions or comments can be implemented to enhance communication. Also, including users in the design or testing process will enhance their acceptance and use of the script, as they can familiarise themselves while it is being developed. Afterwards, they can take their gained knowledge to their own organisation. Lastly, the interviewees prefer to use automation as a method to advise or inform them. For example, an automation script can indicate if an issue occurs between your spatial and temporal resolution. After this, the modeller can make a modelling decision about what to change. Or, as an interviewee indicated, a modeller can be given advise by the automation tool as to what the best temporal resolution is.

Because the efforts to centralise automation scripts at the inter-organisational level are still ongoing, not all interview codes based on the theoretical concepts were applicable in the content analysis. Some interview codes, such as 'Redefining the modellers role' or 'Different skills / knowledge obtained by the modeller', were not used because not enough time has passed to observe these changes. The same is the case for the interview codes related to the accuracy of the automation's perceived trustworthiness. For these interview codes, it might also matter that some modellers were involved in the development or testing of automation scripts, which results in a different first impression.

3.3.3 | Interaction between Water Authorities and Consulting Companies

The interviews showed that the roles of the water authorities, consulting companies and research institutes propagate into the modeller's perceptions on automation, resulting in differences between the water authorities and consulting companies. A modeller at a water authority more easily accepts automation in any of the steps executed by a consulting company since they already transferred this modelling decision to an external partner. For example, if a water authority used data provided by a research institute, the pre-processing of these data is already done by the research institute. Then, as an interviewee stated, "*I just assume that this [automation of pre-processing] already happens at the research institute. Again another assumption, I don't know what exactly happens there.*". Even though the modelling step is executed by an external partner, it is still desirable to have transparency and reproducibility. Automation at a (Inter-)organisational level can accommodate this.

Often, similar organisations execute specific parts of the modelling process, e.g. consulting companies tend to execute the calibration and model schematisation steps. Because each organisation is responsible for different modelling steps, it is generally expected that automating certain steps is the responsibility of the organisation that executes it. For instance, automating data pre-processing is the perceived responsibility of the water authorities or research organisations, according to an interviewee at a consulting company.

The general division of these modelling steps between water authority, consulting company and research institute is due to the availability of resources, such as computational power, expertise, staff and funding. Consulting companies generally have more resources than water authorities, especially computational power, but also experience and expertise, to execute calibration, model set-up and validation. Because of the resources needed to automate, water authorities, consulting companies and research institutes need to consider if automation is worthwhile.

3.4 | Reflections and implications

The expectation is that standardisation will result in higher consistency between modelling studies and more transparency. But, achieving these outcomes depends on its documentation, evaluation and maintenance. Consistency and transparency in the modelling process are considered important for model results to be useful in a decision-making context (Watson, 2005). If transparency and consistency are achieved, then decision makers have an improved foundation for their decisions, even though the model results are only one part of their considerations (Calder et al., 2018; Watson, 2005). However, our results indicate that the interaction between automation and its usage might affect the transparency and consistency achieved through standardisation for quality assurance.

There are also other approaches to achieve more transparency, for instance through creating records of engagement and decision making (Cockerill et al., 2019) or audits. All these procedures are meant to achieve quality assurance for decision makers – the persons making decisions informed by model results (in this case study e.g. the local water manager).

Automation, like other technologies, shapes its use, and vice versa (Melsen et al., 2018b). The usage of automation can change over time and the distance between programmer and modeller can increase. This increased distance can produce barriers to communication and examination or questioning of the assumptions in the automation, making the need for the development process and final script to be documented essential (Calder et al., 2018; Wang et al., 2023). While modellers familiarise themselves with any automation script, they might become less critical of the automation and its capabilities and limitations (Parasuraman et al., 2000). This can result in the modellers blindly accepting the automation, of which an interviewee gave an example, “*So basically, it is automated: as a modeller you can change the table and put in other values. However, in reality, that rarely happens, I think. So, at some point, someone came up with these numbers and over time we become a bit blind to this, sometimes too blind.*". The blindness can affect the transparency, because the modellers no longer understand and evaluate what the results are based on. In general, the loss of understanding in the automation can lead to surprising results (Carver & Turoff, 2007; Parasuraman et al., 2000), leaving the modeller the difficult task to figure out what these results mean.

How automation developed at Organisational and Inter-organisational level is used, will determine if it can actually serve as a standardisation method for quality assurance – its usage shapes the automation. Calder et al. (2018) mention that ‘*models might be used for purposes beyond those for which they were originally designed*’ and this also applies to automation. Furthermore, automation scripts can be taken locally. Here, they can be developed further, while no feedback is supplied to the central automation. This practice could potentially create many versions of a similar automation script again, transferring the automation back to Individual or Team automation level (see Figure 3.3). Also, automation scripts can be used in such a way that the modeller adapts it to get the hydrological response they would expect. This was mentioned by some interviewees, with for example an interviewee stating, “*Sometimes, you are fumbling about, misusing certain features in order to get the system response that you want.*", raising the question: ‘Who is making the modelling decision, the programmer of the script or the modeller?’

The (mis)usage of the automation can potentially negate its intended transparency and consistency at an organisational or inter-organisational automation level, affecting the standardisation potential of automation. For maintaining the quality in models or technologies, Calder et al. (2018) suggest that a proper review system should be set up. At some point, this review might result in needing to retire some of the technology developed. Part of a review system is appointing long-term ownership to give someone the responsibility to keep up the review. While for models the owner is usually quite clear, this is more difficult for automation scripts beyond the individual or team level. To determine ownership is especially important for automation because automation potentially impedes the questioning of assumptions more. The power relations between auditor and client is another aspect influencing the questioning of automation scripts (McCracken et al., 2008; Carlisle et al., 2023). Multiple quality assurance and control procedures exist each with their own (dis)advantages: e.g. code review (Bacchelli & Bird, 2013; Pascarella et al., 2018), documentation (Kajko-Mattsson, 2005; Parnas, 2010) and automated testing (Bartram & Bayliss, 1984; Winkler et al., 2010). Besides improved quality, additional benefits can occur such as knowledge transfer or increasing team awareness (Bacchelli & Bird, 2013).

Organisations internally apply a mixture of these quality assurance and control procedures to a certain extent already, as some of the interviewees indicated. Still, as long as no errors occur, the script, data or technology are often accepted as is. If errors do occur, automation requires that the whole process is evaluated more carefully. According to our interviewees, in some organisations, automation scripts were developed relatively recently and the evaluation has not been set up (yet) at the time of the interviews. The allocation of time and resources are important intervening factors, just as the perceived long-term usefulness of the script or tool in question. Inter-organisationally, setting up these procedures takes even more effort. The HYDROLIB project currently has determined ownership for each component of their project and has set up a review system. Generally, regular reviews of automation at the inter-organisational automation level would ensure fit for purposeness and clear documentation, for which Hamilton et al. (2022) and Calder et al. (2018) have, respectively, emphasised. Regular reviews could potentially also help decrease unintended blind spots in the model adaptation. Only when these requirements are met will automation at the inter-organisational level contribute to standardisation as a means to achieve quality assurance, and with that, to consistency, transparency and reproducibility.

3.5 | Conclusion

In this study, we explored the modeller's perspective on automation of the model setup using a case study of hydrodynamic modelling in the Netherlands. We conducted fourteen interviews with modellers at water authorities and consulting companies. Subsequently, we carried out a deductive (based on Pagano et al. (2016)) and inductive content analysis on the transcripts. The analysis resulted in an overview of the different social levels of automation practice at which automation is applied uniformly (Individual, Team, Organisational and Inter-organisational); insights into the current extent of automation and the preferred practices in automation by the modeller; and the interaction between third-parties that conduct part of the modelling work in light of automation. The automation levels cover who created and used the automation script: for the individual, team, organisational and inter-organisational level, this is, respectively, one modeller, within a team of modellers, within one organisation and between different organisations.

Automation of parts of the modelling process has numerous advantages: efficiency, reduction of human error, transparency, reproducibility and consistency between different modellers. These advantages appear at different social levels of automation practice. Efficiency and reduction of human error show at all automation levels, while transparency, consistency and quality assurance for decision makers can be achieved when automation is implemented at Organisational and Inter-organisational automation level. Automation has to be implemented carefully at these (Inter-)organisational automation levels, as modellers might not be inclined to use it if they do no trust it or have a different view on how modelling should be done. Furthermore, automation as a standardisation procedure requires among others documentation, evaluation and maintenance in order for the advantages of standardisation to surface.

Even if the automation is used and has all the requirements for a standardisation procedure, it is crucial to keep testing ones own understanding and understanding the automation (at any automation level). Automation at organisational and inter-organisational levels might include the risk of decreasing transparency, i.e. not being aware of underlying modelling assumptions, or (mis)using the automation in a such a manner that it still aligns with the modeller's perception of the hydrological system. Any project targeting automation at the (Inter-)organisational automation level has to find a balance between creating transparency and consistency with their automation tools and safeguarding the modeller's trust and proper use of their automation tools. This requires setting up a review system and determining who has long-term ownership of the automation tool. Automating the modelling procedure is a fine line between giving and taking power to modellers and programmers, managing the risks and benefits of automation.

4

Comparing standardisation approaches

This chapter is based on:

Remmers, J.O.E., Guillaume J.H.A., Melsen, L.A. (2024). A True North: how does it evolve in hydrological modelling? A comparative study of modelling standardisation approaches in Australia and the Netherlands. *Environmental Modelling & Software (in review)*.

Abstract

In hydrological modelling, differences in results arise from the individual conducting the study, leading to inter-modeller variability. Each step in the process involves numerous methodological choices, shaping a narrative conveyed by inherently non-neutral results. However, these results are frequently used to support decision making, necessitating a clear justification of how model development occurred. Standardisation is deemed to enhance modelling credibility, legitimacy, reproducibility and transparency, though challenges exist. Since modellers must adopt standardisation approaches in their practice, understanding modellers' perceptions of them is important. This study explored two approaches: a Dutch modelling guideline and an Australian common modelling framework. Interview transcripts were analysed using inductive content analysis. Both cases show that building consensus, which is necessary for standardisation, often involves power relations, which can marginalise certain perspectives. Since standardisation is not neutral, it requires careful development, implementation and reflection on what is standardised, who determines it and who ultimately benefits.

"There were, in fact, so many things to be attended to,
so many people to be pleased,
that there did seem as little chance of a decision."

—Jane Austen, *Mansfield Park* (1814)

4.1 | Introduction

In radiology, CT- and MRI-scans are used to delineate anatomical structures in our bodies, such as tumours. A medical expert estimates the size of these structures based on the scans. However, this process can yield different results dependent on which expert made the estimation, a phenomenon known as inter-observer variability (Joskowicz et al., 2019; Sørensen et al., 1993; Zir et al., 1976). To address the inter-observer variability issue, various methods have been developed and tested in radiology. These methods include the implementation of standardised protocols and the application of artificial intelligence (Sharp et al., 2014; Wong et al., 2020). Inter-observer variability is not unique to radiology; similar inter-observer variability has been discerned in other scientific disciplines, such as archaeology (Beck & Jones, 1989; Gnaden & Holdaway, 2000), ecology (Goodenough et al., 2020; Morrison, 2016) and meteorology (Daly et al., 2007). This variability can lead to different research outcomes, dependent on who conducted the research.

In hydrological modelling, differences in research results arise from the individual who conducted the modelling study, leading to inter-modeller variability - a specific instance of inter-observer variability. Numerous methodological choices can be made at each step of the modelling process, for instance based on expert judgment or thorough testing (Melsen, 2022). Holländer et al. (2009) asked ten modellers to conduct the same modelling study. They found that the modelling results varied widely between modellers and that personal judgement played a key role throughout the modelling process. Additionally, the modelling process is substantially influenced by the team, the organisation, and other social, technical and political considerations (**Chapter 2**, Melsen, 2022; Saltelli & Di Fiore, 2023). Each choice shapes the narrative conveyed by the model results (Lahtinen et al., 2017). As such, model results are never neutral (Saltelli et al., 2020). Despite this, the model results are still used to support decision making.

The use of model results in decision support requires transferring information across different disciplines, bridging technical details with decision making (Cash et al., 2002; Hamilton et al., 2022). Different actors will have different expectations and interpretations of the model process and results (Cash et al., 2002; van Voorn et al., 2016). Several factors are important to consider in transferring models results to decision-makers: credibility, legitimacy, reproducibility, salience and transparency (Cash et al., 2002; Hamilton et al., 2022; Schwarz et al., 2020; van Voorn et al., 2016). Credibility relates to basing the model on sound scientific knowledge (van Voorn et al., 2016). Legitimacy relates to representing the views, values and concerns of involved stakeholders fairly within a model (van Voorn et al., 2016). Reproducibility relates to whether a modelling study can be recreated by another modeller, team or organisation (Gundersen, 2021; Schwarz et al., 2020). Ideally, this would result in matching outcomes without unforeseen errors or discrepancies. Salience relates to the relevance of the model results for stakeholders (Cash et al., 2002; van Voorn et al., 2016). Transparency relates to being open about the modelling process: assumptions, uncertainties and values underlying the modelling process should be interpretable and understandable. There are several approaches to improve these concepts in the modelling process.

One way to enhance credibility, legitimacy, reproducibility, salience and transparency of the modelling process is through standardisation. As defined in **Chapter 1**), standardisation is defined as applying the same modelling principles, or the same modelling process, to harmon the modelling process. This would lead to commonly agreed upon norms and values within modelling (Goulden et al., 2019). Standardisation can also facilitate the quality assurance of the standardised process itself (Seth, 1991). Over time, a standard approachcan progress and evolve. Standardisation in modelling can be implemented, for instance, through creating a set of guidelines for general, voluntary use, through prescribing certain guidelines for the modelling process by top-down initiatives, or through automating the modelling process (**Chapter 3**). Each standardisation method comes with its own advantages and disadvantages. To study this, we explored and compared two standardisation-cases: a modelling guideline in The Netherlands and development of a common modelling framework in Australia.

A comprehensive modelling guideline has been developed in the Netherlands (van Waveren et al., 1999). This guideline was developed in response to a felt need from the community itself, as they observed the wild west in ongoing modelling practices which lead to errors and inconsistencies. The guideline is broad with the objective to assist Dutch hydrological modellers in supporting water management decision making and to improve modelling quality in the Netherlands. Similar guidelines have been created in other disciplines as well, such as Philips et al. (2006) for health technology assessments, Caro et al. (2012) for updates on that and Jones et al. (2020) for environmental modelling. In general, guidelines tend to be developed to address a gap, for example differences in terminology or process understanding, between the people creating or using models and the people being informed by the model results (Refsgaard & Henriksen, 2004). However, adhering to these guidelines requires an investment, for which the necessary resources are sometimes lacking (Jones et al., 2020). As such, the development of a guideline is only a first step, the next challenge is to have the guideline being used in practice.

Standardisation can also be approached through top-down initiatives, such as the development of a common (national) modelling framework. The Australian government has initiated the development of eWater Source, a hydrological modelling framework (Parliament of Australia, SA). The reason for this development was a severe drought (the Millennium drought), which required water managers to collaborate across state borders. Other examples of national modelling frameworks are MAGPIE, which was created to evaluate nitrate losses in the UK (Lord & Anthony, 2000), NHI, which is a national hydrological modelling framework in the Netherlands (Hoogewoud et al., 2013; de Lange et al., 2015) and the DK-model, which is a national hydrological model for Denmark (Henriksen et al., 2003). NHI and the DK-model were initiated by independent research institutes. For MAGPIE, the UK's national government had a vested interest in and necessity for the development of this model framework, because the Ministry of Agriculture needed to estimate nitrate exports.

As the hydrological field progresses and water-management issues become increasingly urgent, countries and organisations adopt different standardisation approaches for decision-support modelling. Standardisation is the result of a social process during its development and adoption (Alder & Wise, 1995; Jimenez, 2022). Since, after all, the modellers are the ones who have to accept and adopt these methods in their working practice, it is relevant to understand how different standardisation approaches are perceived by the modellers themselves. With this study, comparing two standardisation approaches in Australia and the Netherlands, we aim to discern what effects standardisation approaches have on modelling practice.

4.2 | Methodology

4.2.1 | Case Studies

Two contrasting case studies were selected: modelling the Murray-Darling Basin in Australia and modelling for water management in the Netherlands. They differ in water management issue, governance arrangements and standardisation approach. Both case studies still have the same intention: harmonising the modelling process.

Australia — In Australia, water management is focused on addressing water quantity and quality limitations. Water management is executed at state level, including how water is allocated. Australia implemented a water allocation system to manage its water resources. However, there has been an over-allocation of water entitlements in the Murray-Darling Basin (MDB). Therefore, in the early 90s, a cap was set on the water entitlements (Parliament of Australia, SA). Each state is responsible for granting and checking compliance of the water entitlements (Grafton & Horne, 2014). To support water management decisions, each state has developed and implemented their own strategy for hydrological modelling, utilising rules-based and linear-programming based modelling (MDBA, 2012; Connell & Grafton, 2011; Yang et al., 2017). However, some basins cover multiple states, most notably the Murray-Darling Basin (MDB, Yang et al., 2017). In this study, we focus on the Southern part of this basin.

Initially, the existing sub-catchment models, developed in the so called legacy models – the previous model software developed at state level, used for decades –, were coupled to tackle basin-wide questions. This modelling framework was the Integrated River System Model Framework (IRSMF, MDBA, SA), which was created in 2008. In cooperation with the states and the Murray-Darling Basin Authority (MDBA), a new modelling framework was developed to cover the whole basin. The national government facilitated this process. This development was instigated due to the Millennium Drought (2001 – 2009). During this event, water resources requirements were not met for environmental flows and human needs, which also includes the water entitlements (Kirby et al., 2014). Because the standardisation, this basin-wide modelling framework, starts at the government level, it is a more top-down initiative of standardisation. .

A new model framework, eWater Source, was developed by a consortium between 2005 and 2012 (eWater Source Group, SA). The responsibility for maintenance and funding of the framework was handed to the eWater Group, a newly formed company owned by the federal and state governments. The states committed to replace their sub-catchment models built in their legacy model with models built in eWater Source (Hart, 2016). However, at the time of the interviews, not all sub-catchment models had been transitioned to the eWater Source framework yet, and those that were still had to be stitched together manually.

The Commonwealth has designated additional funds (66 million AUD) to push the transitioning to the Source modelling framework in the MDB further. This is the Integrated River Modelling Uplift (IRMU) programme, which will run for 4 years. The funding was acquired at the end of 2021 and the programme will run until mid-2026, but work had already begun in 2020. Additionally, the aim is to further integrate and standardise the 24 subcatchment models on a cloud-based framework, minimising manual interventions. The intention is to have all 24 subcatchment models in eWater Source. Part of IRMU is also dedicated to make the models fit for future questions and decisions (MDBA, SA).

The Netherlands — In The Netherlands, which generally is not water limited, water management focuses on flood protection, maintenance of the water ways for transport, and drought or water surplus prevention (Dutch Water Authorities, SA). Modelling is used to support decisions in any of these situations. decision-support modelling for water management is mainly the responsibility of the national and local water authorities and supported by consulting companies. 21 local water authorities execute the water management in their area (Government of the Netherlands, SA). The governmental agency Rijkswaterstaat functions at the national level (Government of the Netherlands, SA; Vukovic, 2022). Dutch consulting companies assist the water authorities. This support can cover the execution of the whole modelling process or knowledge provision to water authorities such that they are able to set-up and execute models themselves. Research institutes, such as Deltares, co-develop and maintain model software, ensuring its quality. A more elaborate description is provided in **Section 2.2.1**.

The guideline 'Good modelling practice: Vloeiend modelleren in het waterbeheer', published in 1999, was written to initiate more standardisation in the Dutch hydrological modelling community (van Waveren et al., 1999). This initiative was led by individuals from various organisation, most notably Rijkswaterstaat. Dutch water authorities and consulting companies had no obligation to incorporate the guideline. Therefore, guidelines can be seen as a more bottom-up approach. Based on the interviews, after publication, the uptake of the guideline in day-to-day usage was less than anticipated and decreased over the years.

After the publication of the guideline, other initiatives were developed, both at the national and international level. In the Netherlands, an attempt was made to formalise the guideline through NEN-norms, Dutch standards specifying quality criteria, similar to the ISO-standards (van der Molen et al., 2002; Nederlandse Rijksoverheid, SA). Nationally, a common modelling platform for ground and surface water modelling, the Nationaal Hydrologisch Instrumentarium (NHI), was developed (Kukuric, 2011). This development started in 2005 to provide decision support for nationally and regionally relevant hydrological questions. The NHI is being updated regularly (Hoogewoud et al., 2013; de Lange et al., 2015). Internationally, the guideline was followed up in the EU-project HarmoniQua (Kassahun et al., 2005; Scholten et al., 2004). The outcome of this project was a computer-based toolbox and a digitised guideline to provide support for quality assurance in decision-support hydrological modelling.

4.2.2 | Interviewee selection and interviews

To acquire the modeller's perspectives on both standardisation methods, we conducted semi-structured interviews with modellers and people involved in the development of the standardisation approach in Australia and The Netherlands. The interviewees were mainly selected through convenience and snowball sampling. The first interviewees were selected through reaching out to suitable people in the authors contacts or projects the authors were involved in (convenience sampling). At the end of each interview, we asked the interviewee if they had any recommendations for other potential interviewee (snowball sampling, Emerson, 2015). Some interviewees were selected through purposive sampling – selective sampling to include interviewees who are experts for either of the case studies. These sampling methods are effective to analyse a certain case study within one geographical area including knowledgeable experts, which our case studies are. All interviews were recorded and transcribed with the interviewee's informed consent with ethics application 2023-013 granted by WUR Research Ethics Committee.

Australia — We conducted twenty semi-structured interviews with users and developers of the standardisation approach in the MDB. These interviews took place in June and July 2023. All interviews were held in English and lasted between 30 minutes and 2 hours. The interviews covered different parts of the standardisation initiative in Australia on eWater Source. The general interview guide has been included in Appendix of Chapter 4A. We adapted the guide depending on the interviewee and their exact expertise. In this study, we refer to these interviews as AUS1 – 20. We interviewed modellers who work or have worked at governmental state agencies or the MDBA, as well as some developers involved in creating the eWater Source framework. Additionally, we interviewed individuals from government agencies, consulting companies and independent experts involved with the modelling, water management or reviewing within the MDB during the development of or transition to eWater Source.

Netherlands — We conducted fourteen semi-structured interviews between September and December 2021 with hydrodynamic modellers, which were also used in **Chapters 2 and 3**. At the time, nine interviewees worked at six different water authorities and five interviewees at four different consulting companies. Two additional interviews were held in the spring of 2023 with developers of the guideline. All were held in Dutch, except one, which was in English. All interviews lasted between 1 hour 15 minutes and 2 hours. The interviews with the modellers entailed some questions about standardisation within The Netherlands. The interview guide has been included in Appendix General. The interviews with the developers concerned the development of the guideline and were specifically tailored to each individual interviewee. In Appendix of Chapter 4A, a general interview guide is provided with example questions. In this chapter, we refer to these interviews as NL1 – 16.

4.2.3 | Content Analysis

The transcripts were analysed in ATLAS.ti, version 9 (ATLAS.ti Scientific Software Development GmbH, 2022). We performed an inductive content analysis, which entails analysing the full interviews and assigning so-called codes to excerpts of the interview transcripts. In this chapter, we determined the codes as we analysed the interviews. First, topical codes were inserted to categorise which parts of the interviews were related to the standardisation approaches. For the transcripts from interviews with Dutch modellers, we only coded the parts regarding standardisation. The other interviews were coded completely. After the topical codes, we coded the interviews inductively, which resulted in 138 inductive codes. These codes were grouped in seven axial codes derived inductively: Development, Application, Advantages, Disadvantages, Improvements, Risks and Updates. Development relates to the creation of the standardisation approach and its intended purpose. Application relates to how the standardisation approach is employed. Advantages and Disadvantages relate to the actual outcomes of its application. Improvements relate to potential ways to enhance the standardisation approach. Risks relate to potential pitfalls the standardisation approach might have. Updates relate to any revisions or continuations of the standardisation approach. The inductive codes are included in Appendix of Chapter 4B. Results are presented following the axial codes, first for each case study individually and afterwards in comparison.

4.3 | Results and Interpretation

4.3.1 | Case study Australia: the interviewees' perspectives on the uptake of eWater Source

Development and Application — The main motivation for standardisation efforts in Australian water management modelling was a water crisis (Interviews AUS 2, 3, 6, 8, 14 and 20). The Millennium Drought (2001 – 2009) was one of the instigators to initiate the development of the Source modelling framework, because it was no longer tenable to model the subcatchments without incorporating more communication and consistency between these subcatchment models. Consistency increases the ease of use when linking the sub-catchment models for full-basin studies. Additionally, it can create transparency between model studies, modellers and stakeholders. This transparency is needed for the model results that support policy decisions (AUS 2, 4, 10, 12, 16 and 17).

Consensus, through negotiation, is necessary when developing and implementing a common modelling framework. For the modelling practice in the MDB, policy is a major factor, since each state has their own water policy rules that need to be incorporated in the modelling approach. Consensus involved accommodating different policy options and modelling approaches within the framework, as well as making an agreement to transition (AUS 4, 6, 8, 11, 13, 14, 16, 17 and 20). Negotiation was needed to establish a timeline and allocate resources. Although the transition is slow, all states are moving towards or have finished the transition.

Since the models are used for policy and some numbers are directly incorporated in legislation, the implementation of a new modelling framework is executed carefully. Transitioning tends to follow the policy updates, which are commonly on a ten-year cycle. Furthermore, the transition is resources-limited, with available funding and people stretched between current modelling tasks and the transition to the Source modelling framework.

In order to build trust among stakeholders, several interviewees (AUS 2, 4, 6, 8, 9, 14, 15, 17, 18 and 20) indicated that it is crucial to be able to explain the differences in results when transitioning to a new modelling framework. This might involve reconstructing the legacy model in the new framework and then gradually modifying it (AUS 4, 12 and 20). The models used to support policy generally are independently reviewed, both the legacy and Source models. These reviews are another way to increase the trust in the models in a policy context.

Throughout developing and transitioning to eWater Source, a community was built through, for example, involving modellers in the development of eWater Source, providing training and creating a general guideline. The community building included various stakeholders across the MDB and within each state, including modellers and developers. One aspect of community building is the Practice Notes, voluntary guidelines for Source users aimed at create consistency in modelling practices. The Practice Notes are based on the first user cases of Source. The interviewees were generally aware of the existence of these Practice Notes. However, they are not used on a daily basis (AUS 4, 9, 11, 12 and 14).

Advantages and Disadvantages — Differences in terminology between states hinder the development and adoption of one modelling platform (AUS 1, 4, 8, 11, 14, 16 and 17). This complicates communication between modellers and raises the question which terminology to use in a common platform. By choosing one terminology, someone's, an organisation's or state's preference has become the default. Of course, definitions can be explained in documentation, but, documentation is not necessarily read.

The existing different approaches between states complicate the development and transition to Source. Being able to use their own approaches within the framework was a demand from most stakeholders in order to commit to the transition to Source. Generally, the Source framework has accommodated the different existing approaches (AUS 2, 3, 4, 7, 8, 9, 11, 14, 15, 16, 17 and 20). Because of this, the Source modelling framework has many varying features and can be quite computationally demanding (AUS 5, 9, 13, 14, 16, 18 and 20). In essence, this means that transparency is increased with the implementation in Source in the sense that everyone is more familiar with the framework applied, but not necessarily with the approach taken within this framework.

Improvements, Risks and Updates — Transitioning to one modelling framework can pose some risks and opportunities. These can guide the progression of the standardisation approach.

One potential risk is the future maintenance of the framework and the resources this requires (AUS 2, 6, 8, 13, 15, 16 and 17). As the framework is specifically built for the Australian context, so far it is mainly used in Australia. Because of this, there are concerns with regard to its long-term financial viability, which is the responsibility of the eWater Group. Currently, the maintenance of the framework requires structural input from the Commonwealth and different states.

The complexity of the framework poses another risk. This complexity means that the upkeep costs are substantial (AUS 6, 8, 13, 15, 16, 17 and 20). Hence, a critical look at whether all features are necessary might reduce future maintenance costs (AUS 15). However, this complexity was also necessary to get everyone on board.

The longevity of the framework is another concern. The high investment costs made combined with the uncertainty of the sustainability of the Source modelling framework leaves the Australian modelling community open for the risk of a lock-in effect – a system is dependent on one option while not having the possibility to switch without substantial costs. Two interviewees (AUS 17 and 18) expressed their concern for the situation in which Source is discontinued, which will result in massive costs to put something else in place. Other interviewees were concerned that "*Source might fade before it gets finished, but hopefully it won't*" (AUS14) or were wondering "*when Source* *would be seen as a legacy model*" (AUS15). Contrary to this, another interviewee mentioned that "*despite the frustrations in the early days, people still kept using it says something about the dedication of many of those users*" (AUS11).

The dedication to keep working with the Source modelling framework can be used in the IRMU programme, for which the interviewees recognised the aim to cooperatively improve the technology (further development of Source and the development of a cloud-based framework). Identified advantages of this include more collaboration (AUS 2, 4, 6 and 16), transparency (AUS 2, 4, 6, 7, 9, 13 and 16), efficiency (AUS 1, 2, 3, 6, 7, 8, 9, 12, 13 and 15) and consistency (AUS 2, 4, 5, 6, 11, 12, 13, 15, 16, 17 and 20). Furthermore, it is mentioned that this programme provides the funding necessary to accomplish the improvements (AUS 2, 15 and 16). Within this programme, it is difficult to align different priorities (AUS 2, 7, 8 and 9), to determine long-term ownership (AUS 2, 13, 14 and 16) and to provide continuous funding (AUS 2, 6, 8, 13, 15, 16 and 17). The timeline of the programme is also quite ambitious, but still an important step in furthering the transition to Source, even if it might not accomplish every goal (AUS 13, 16 and 17). Communication between stakeholders will be important to keep evolving the framework and setting out the direction it will take. Part of this will also be to determine which processes need to be represented in the modelling and how the science can progress.

4.3.2 | Case study The Netherlands: the interviewees' perspectives on the uptake of the guideline

Development and Application — The development of the Good Modelling Practice guideline was mainly informed by the motivation to improve model quality and transparency in the Netherlands (NL 15). This motivation was shared by multiple modellers and culminated in Rijkswaterstaat leading and creating a collaboration between multiple organisations, both from industry and public partners (NL 15 and 16, van Waveren et al., 1999). Even more people and organisations were involved during the testing phase and in an advisory committee.

The idea behind involving many different organisations and people was to create a community. This provides a support base for the usage of the guideline (NL 15). Implementation hinged on voluntary usage. The intention was that as part of the community, commissioners would ask for a modelling study to be done according to the guideline and that any organisation executing a modelling study would provide a framework internally to use the guideline (NL 9, 15 and 16).

The voluntary implementation uncovered that there was less consensus than it seemed during the development. From the interviews it appeared that the uptake of the guideline was less than anticipated (NL 1, 2, 3, 5, 6, 7, 9, 10, 11, 13, 15 and 16). Currently, modellers barely use the guideline actively: for example "*I think that modellers have what is important in the back of their minds, but that no one has the guideline next to them. At least, I don't know anyone who does.*" (NL 5). Interviewees think the guideline have a high level of generality, which hampers the active usage (NL 1, 2, 3, 13 and 16). The interviewees indicated that they rather use it as a general framework in the back of their heads (NL 5, 8 and 13). Other usages of the guideline include basing the organisation's modelling vision on the guideline (NL 4, 10 and 13) or giving the guideline as a reference to new or less experienced modellers (NL 3, 6, 7, 8, 9, 10 and 14). With these more indirect usages, the guideline have shaped modelling practice and thinking by making steps explicit that have become institutionalised.

Advantages and Disadvantages — The Dutch guideline has several disadvantages and advantages. Disadvantages of the guideline encompass the site-specificity of modelling studies (NL 2, 4, 11 and 13, e.g. "*every modelling study is still some customisation*"), the existing different practices between modellers and organisations (NL 2, 3, 6, 10, 13, 15 and 16, e.g. "*people preferably want to use what they're used to*"), the length of the guideline (NL 3, 4, 7 and 10, e.g. "*I can imagine that the longer such a guideline is, the less appealing it will be*"), and the high effort in using the guideline (NL 3, 4, 5, 13, 14 and 16, e.g. "*You didn't want to have too much administration, so you could only spend half your time on modelling*"). These disadvantages can be reasons for the low direct usage of the guideline. Despite these disadvantages, interviewees (NL 3, 6, 7, 8, 9, 10, 13 and 14) did recognise that the guideline can be beneficial: "*of course, it also undoubtedly contains some very good and logical things*".

The advantages mentioned in the interviews include making better choices, sharing of problems and information (i.e. one form of transparency). For making better choices, the guideline can aid in providing a reference for the modellers (NL 3, 6, 7, 8, 9, 10 and 14), reducing human errors and improving reproducibility. The development of the guideline created a community (NL 15 and 16), in which problems and information are shared. In this way, it provides a knowledge base for modellers. Guidelines can also reduce inter-modeller variability, i.e. create consistency, through facilitating collaboration and establishing a common language (NL 15 and 16). This can potentially increase the reproducibility of modelling studies and transferability of modelling skills.

Improvements, Risks and Updates — Based on the guideline, there have been a couple of follow-up projects, which show how standardisation approaches can progress. These follow-ups highlight the importance of consensus and the difficulty of overcoming different approaches. An attempt was made to formalise the voluntary guidelines into NEN-norms – Dutch standards similar to ISO-standards (Hoogewoud et al., 2013; de Lange et al., 2015). However, these were never accepted. The different parties were not able to reach a consensus on what the norms should prescribe in terms of modelling practice, due to different modelling approaches. Within the European HarmoniQua project, multiple international organisations were involved. In this project, a tool was developed. The tool was used, though mainly by the people involved in the project. Within the Netherlands, HarmoniQua was not really taken up beyond educational purposes initiated by one of the project members. The maintenance was also done by the people involved in its development, resulting in its disuse after the developers' retirement or move to different organisations.

Following its publication, the guideline was updated once or twice in the beginning, showing the guideline's adaptability. Initially, a printed version of the guideline in a multomap was distributed to everyone, ensuring easy adaptability (NL 15, 16). The users could send in suggestions. Multiple suggestions were incorporated at the same time. The changed pages of the guideline would be sent to everyone, which they could easily swap in their multomap. Another source of suggestions came from application in education for a few years (NL 16). As a voluntary standardisation approach, the guideline potentially suffered from lack of ownership and resources. STOWA – the Dutch centre of expertise of the water authorities – is the owner of the guideline. However, STOWA is a demand-driven organisation, meaning they do not have the capacity (funding and humanpower) to actually maintain the guideline and promote its usage. Over time, this decreased the incentive to keep using it. This reduced the possibilities for updates and promotion for its usage. Even though HarmoniQua was a follow-up project, it did not provide funding for the Dutch guideline as its focus was on an international tool.

The guideline was not adopted actively, decreasing its longevity, although the guideline is informing vision and used in the back of modellers' minds. In this form, the guideline has evolved. Interviewees have given further suggestions for evolving the guideline: create a shorter version or summary (NL 3, 7, 10), change title to reflect its contents better (NL 4), integrate the guideline with modelling tools (NL 2, 9, 14), or create a digital version of the guideline (NL 4, 6, 15). Two interviewees (NL 6, 7) indicated that is was mainly important to have an easily searchable format, regardless of its format.

4.3.3 | Commonalities and differences

In this section, we will explore commonalities and differences in both standardisation approaches. The main differences in the development and implementation of the standardisation approaches are in the initiators and its usage. The eWater Source is modelling software, rather than guidelines. The different stakeholders agreed to transition to this software. This approach is more of a top-down approach due to the Australian Commonwealth being the initiator. The Dutch guideline is a voluntary approach, using soft power. The guideline is barely actively used in water management modelling. This approach can be classified more as a bottom-up approach due to the initiative coming from the modellers themselves.

Development and Application — The motivations for developing a standardisation approach is different in both case studies. In the Netherlands, the emphasis is on improving model quality as this is important for the guideline's initiators – the modellers. In Australia, the emphasis is on creating legitimacy for policy and legislation as this is valued by the eWater Source's initiators – policy-makers.

Despite their different goals and initiators, we recognise that a consensus needs to be reached in both instances, which includes building a community for implementation. For the Dutch guideline, a consensus was reached and a community was build. However, in practice, this consensus did not seem to reach its intended use, since modellers reverted to their own modelling practices. Another example of the importance of consensus in standardisation are the NEN-norms, which would give a more formal status to the guidelines. These were not developed due to the absence of a consensus. For the Australian modelling framework, a consensus was also reached in an agreement to transition to the Source modelling framework. After this consensus, the community has been build around the development and transition of the framework. Partly because of this prior binding agreement, the different stakeholders keep using and transitioning to the framework. Grimm et al. (2014) also recognise that good modelling practice has long been known, but the problem is to get it established. Our results confirm this, demonstrating that, without a strong concerted effort, it is difficult to formalise and standardise modelling approaches.

The differences in top-down and voluntary approaches leads to different adoption of the standardisation approach. Implementing the Dutch guideline, a voluntary approach, proved to be quite difficult. Implementing the top-down Australian modelling framework also had its difficulties, but the transition is still happening. Goulden et al. (2019) discern these differences between voluntary and top-down approaches in standard setting for standards in environmental policy. They found that with voluntary standards, it is more difficult to implement them, compared to standards defined in law. However, with the voluntary approach, soft effects are visible, such as the changes in modelling vision of organisations. As such, one could even argue that the guidelines achieved their goal of improving modelling practices.

Advantages and Disadvantages — Regarding the relative merits of guideline or tools used as a standardisation approach, some differences are perceived. In the Australian case, the approach is expensive in its development and maintenance. Furthermore, despite its usage, there was initially dissatisfaction about the software, which is slowly improving through fixing bugs in eWater Source. In the Dutch case, the standardisation approach is relatively cheap. However, there is a problem with regards to its adoption: the guideline is barely actively used.

For both case studies, the interviewees see value in creating consistency, reproducibility and transparency. These values can enable transferability of modelling studies and skills between organisations, as well as enhance the credibility and legitimacy of model results in a policy setting. Stakeholders can more easily check on what the model results are based. Especially the Australian case study highlights the necessity of legitimate model results. The modelling results are directly impacting water allocation. The justification needed by stakeholders when model results change can be more easily provided when models are transparent, reproducible and consistent.

Despite these added values of standardisation, both case studies show that there are also some aspects that hinder implementation of it or that are perceived as disadvantages. First, multiple interviewees from both cases indicated that one general good modelling practice or modelling platform is not feasible (AUS 5, 10, 18 and 20, NL 5, 11 and 15). For instance, they highlighted the site-specificity of modelling studies (AUS 5, 6, 8, 14, 15, 16, 18 and 20, NL 2, 4, 11 and 13). Within the usage and implementation of eWater Source, Australian water management modelling has moved to a lock-in effect, which potentially impedes the possibilities to implement site-specific methods. However, the current set-up of eWater Source contains many different possibilities, thereby still providing ample flexibility.

Another aspect hindering standardisation's implementation is the different approaches between organisations. Every organisation prefers to use what they are accustomed to (Addor & Melsen, 2019, AUS 5, NL 16). So, negotiation and compromise is needed to reach a consensus on which modelling approach to incorporate in the standardisation approach. However, if this approach does not align with the organisation's previous modelling approach, organisations might revert to using their previous approach. In both case studies we see that multiple stakeholders were involved in the development and multiple views are represented. Nonetheless, the modellers in both cases were (eventually) able to keep using their own approach. The Dutch guideline never reached an official status and became less used over time, while eWater Source became more complex to accommodate all the different modelling approaches within the framework. In this way, both approaches lost some value in their aim to standardise.

The standardisation's loss of value also elucidates the powers that are at play in the implementation of standards. This especially became evident in the Australian case about the maintenance of eWater Source. The maintenance had to be carefully managed as different stakeholders assigned a different urgency level to bug fixing or necessary development. This creates an imbalance in whose problems are dealt with first. In general, negotiation and compromise necessary for developing a standardisation approach involve power relations. These power relations might result in some perspectives becoming marginalised during the standardisation procedure, while other perspectives become formally adopted.

Improvements, Risks and Updates — The longevity and evolution of a standardisation approach is dependent on several factors, as highlighted by the cases. From the Australian case, it can be seen that simplicity, where possible, can be of value for the maintenance and user-friendliness of the modelling software. More complexity tends to increase the maintenance costs and decrease the ease of use of a tool, although this complexity was necessary in the first place to have enough support from the participating modelling institutes.

Both case studies indicate that resources are necessary for continuous maintenance and further development. eWater Source and the Dutch guideline were initially created during a project. This sets a time limit on the funds available. After the project finishes, new funding needs to be acquired. So far, funding has been procured for further development and implementation of eWater Source. For the guideline, no further funding was made available. It is a common issue that less funding is available for maintenance than for development (e.g. in the IT sector (Daniels & Starr, 1997)). One negative consequences of this is the inability to meet the expectations set during development and implementation of the technology. If the technology is well-maintained, it would increase its credibility among the stakeholders (Daniels & Starr, 1997).

The Dutch guideline shows that adaptability of and within a standardisation approach is important. Adaptability provides the opportunity for a standardisation approach to evolve. With new insights, it should be possible to change the standardisation approach (Meijer et al., 2023). As several interviewees have indicated, adaptability should also be possible within a standardisation approach. For example, due to site specificity, a modeller might want to change the specific method. To take site specificity into account, but still keep a similar modelling approach, consistent inconsistency is necessary. The specific methods can change, i.e. be inconsistent, but the modelling approach (e.g. which modelling steps are taken) should be similar, i.e. be consistent.

4.4 | Discussion

4.4.1 | Standardisation from different perspectives: Human, Organisation and Technology

In our study, we have taken an individual perspective on standardisation approaches through conducting interviews with individual modellers and developers. As such, we propagate that within standardisation often the perspective of the modeller is taken (Jones et al., 2020). Other studies have shown that within modelling also other perspectives, such as the team, organisation or inter-organisational network, are important (**Chapter 2**, Babel et al., 2019; Melsen, 2022). Therefore, it is important to also consider these perspectives when evaluating standardisation.

A broader perspective to standardisation has been brought together in the Human, Organisation and Technology (HOT)-fit framework (Fig. 4.1 Yusof et al., 2008). Yusof et al. (2008) developed the HOT-fit framework for analysing the adoption of Health Information Systems. Health Information Systems are tools implemented to standardise healthcare. The three different factors, Human, Organisation and Technology, need to fit with each other in order for the tool to be adopted. Besides the fit between the factors, the factors also impact each other. The standardisation approach needs to show it has Net Benefits, in which the factors Human, Organisation and Technology come together. Although initially developed for Health Information Systems, the framework can also be used for the standardisation approaches in our case.

HUMAN — The Human perspective looks at the individual user of a standardisation approach, which contains two aspects: System Use and User Satisfaction. System Use means how much the System is used by an individual. This aspect can be evaluated by factors including how many times, by whom, how and why it is used. User Sat-

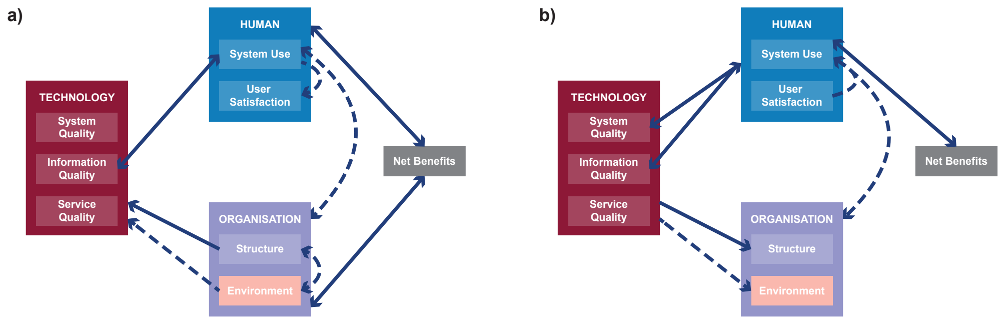

Figure 4.1 | The relationships between the factors, Human, Organisation and Technology, concerning the implementation of standardisation approaches. a) shows the relationships identified from the interviews for the Australian case study and b) presents the relationships identified from the interviews for the Dutch case study. This is adapted from the Human, Organisation, Technology (HOT)-fit framework from Yusof et al. (2008). The solid arrows are present in the HOT-fit framework as describe by Yusof et al. (2008). The dashed arrows are additions we made based on our results.

isfaction means if individuals appreciate the usage of the standardisation approach. Evaluating factors include perceived usefulness, overall satisfaction and enjoyment.

With regards to User Satisfaction and System Use, the two case studies are quite distinct. The adoption of the Dutch guideline follows the HOT-fit framework as described in Yusof et al. (2008) quite closely. The modellers did not experience any Net Benefits from using the guideline actively, especially because the end users did not value this extra effort. Therefore, the User Satisfaction became less, and thus the System Use became less. As the usage decreased, so did the Net Benefits. In the Australian case, the HOT-fit framework as described in Yusof et al. (2008) does not seem to fully explain why the eWater Source framework is adopted, as the original framework does not have a direct link between Human and Organisation. As indicated, the initial System Use of the framework was quite frustrating at times, meaning lower User Satisfaction. The Net Benefits were probably not experienced yet for the individual modeller. So, following the HOT-fit framework, this would mean that usage would go down. However, the usage of the eWater Source framework is still ongoing and increasing. This is due to the Organisation and top-down approach, enforcing the usage of the modelling framework. At this basin-wide level, Net Benefits were probably experienced in the sense of increased collaboration and increased information to base basin-wide policy decisions on. Therefore, we have added a dashed arrow between Organisation and System Use (Fig. 4.1) to the framework to indicate that the Organisation can stimulate the individual usage of the standardisation approach. This can also be described as internalisation and institutionalisation (Babel et al., 2019; Melsen, 2022).

ORGANISATION — The Organisation perspective includes the single organisation (Structure) and the broader organisational perspective (Environment). The Structure relates to, for example, the organisation's strategy, management or leadership. The Environment relates to the broader situatedness of the individual organisation. This can include the government, politics, inter-organisational relationships or financing resources.

The Dutch and Australian case studies highlight different links from and to organisation Structure. The Australian case shows the link between Structure with System Use. Within this link two aspects are clearly visible. First, the capability, e.g. knowledge, to transition to eWater Source is there, which has a positive effect on the System Use. Second, the organisations do not necessarily have the capacity for the transition. They lack resources and humanpower, which negatively impacts the System Use. Furthermore, the organisations influence the technology. The states' different approaches have been incorporated over time, thereby shaping eWater Source.

The Dutch case features how Technology influences the Organisation. Even though the guideline was not used actively at the individual level, it did shape the organisational Structure, for example in an interviewee's organisation they based part of their vision on the guideline: "*A part of our vision is based on it [the guideline]. We would like to follow that, but if we always succeed...*" (NL 4).

Both case studies have a multi-organisational component, which is partly represented by the Environment and the links it has with other aspect of the HOT-fit framework. For instance, the Dutch guideline was intended to be implemented through the Environment: the commissioner of a modelling study should ask for the guideline to be followed. Additionally, it was expected that the organisation structure would provide the capacity to use the guideline. So, this highlights a link between the Organisation and System Use. With the development of eWater Source, the Commonwealth of Australia was an important initiator, indicating a link between Environment and Technology (added as a dotted arrow to Fig. 4.1). One of the interviewees (AUS 6) mentioned that "*in government, politics tend to trump everything including the tool*", indicating that this link between Environment and Technology can be mainly driven by the Environment.

The Environment can also be shaped by the Technology. For example, the Dutch guideline was the precursor of other initiatives in the modelling community, such as the development of a common modelling platform NHI. This changes the inter-organisational relations and collaborative capabilities. Additionally, we added a dotted arrow between Environment and Structure. For example, the lock-in effect due to the development of a standardisation approach, as the Australian case might have, can mean that the Environment requires the Structure of an organisation to change. Similarly, the organisation Structure can also influence the Environment though we did not clearly perceive this in the case studies, because the case studies had an inter-organisational consensus. This is an extension to the internalisation and institutionalisation (Babel et al., 2019; Melsen et al., 2018a).

TECHNOLOGY — The Technology perspective covers the standardisation approach's System Quality, Information Quality and Service Quality. System Quality is defined by the features of the technology, including criteria such as availability, ease of use, flexibility and reliability. Information Quality relates to how users can obtain the necessary information and what information is available. This can be evaluated by legibility, format, conciseness and completeness. Service Quality is about how easy users can receive assistance while using the standardisation approach. This includes quick responsiveness, follow-up service and technical support.

Regarding System Quality, interviewees indicated several options to improve the Dutch guideline, such as reducing its length, providing an online alternative and increasing its searchability. This is an example of how System Use can potentially influence System Quality. These changes can also impact the Information Quality of the guideline, as they influence the format and the conciseness of the standardisation approach. In the Australian case, the development of Practice Notes, which is part of Information Quality, was initiated during the usage of eWater Source (System Use). Having these Practice Notes can also encourage System Use, while becoming informal standards themselves.

Concerning Service Quality, the case studies show several connections. The Dutch case study highlights the importance of ownership of the Technology. The owner of the guideline, STOWA, does not have the capacity to maintain the guideline. Ownership by either an organisation, such as STOWA of the guideline, or an overarching institute, such as the government in Australia, can improve the upkeep of the standardisation approach. On top of that, having a clear owner can facilitate the technical support for users, for example the updates of eWater Source, although, challenges with prioritisation can decrease the Service Quality.

4.4.2 | Will we standardise?

Even though standardisation can establish many advantages, e.g. reproducibility, transparency, consistency and human error reduction, there are challenges. First, one important challenge is representation within the development of a standardisation approach and within the approach itself. A diverse and fair representation of stakeholders is important in the development of standardisation approaches (Goulden et al., 2019). Within standardisation, often the perspective of the modeller is taken (Jones et al., 2020), which leaves out other stakeholders. In our case studies also the modeller's perspective is chosen, but also between modellers micropolitics exist: We recognise differences in representation, which should be diverse to ensure the quality of the standardisation approach (Meijer et al., 2023; Wiarda et al., 2022). For example, the Australian Practice Notes accompanying the Source modelling framework are based on the first user experiences. Although these notes can evolve over time, they can be quite guiding in how the software will be used. Besides, we observed disagreement about setting priorities in the treatment of bugs in the software. For the Dutch guideline, a selection of modellers were involved with formulating the guideline. Although this guideline was discussed in several panels, it is still only this small group that came with the first proposal. So, a certain selection of people can shape a standardisation approach in a certain direction, foreclosing other possibilities. This poses the question whether standardisation is desirable.

Given the advancement of standardisation in health care, we can learn some lessons from insights gained in this field. Within a standardisation approach, it is assumed that one approach is the most desirable, meaning that variation and heterogeneity are consequently undesirable aspects (Wears, 2015). Standardisation gives '*the illusion of the single answer*' (Berg, 1997; Demortain, 2008). This seems to go against the site-specificity that the interviewees highlighted as important. Moreover, choosing one 'desirable' approach can decrease the incentive to keep challenging and evolving that chosen approach. So, the True North needs to be evaluated and adapted when necessary.

Standardisation can also enforce a hierarchy between different perspectives, while there might be no clear reasons to prefer one method over the other (Demortain, 2008; Law, 2004). This difference in hierarchy can also be a reason that people are less inclined to adopt a standardisation approach: the users might not agree with the approach highlighted in the standardisation. This stems from different approaches between people, in our case modellers and organisations. For example, guidelines might be incompatible with ingrained institutionalised standards (Melsen, 2022). To counter the imposed homogeneity and hierarchy from standardisation to some extent calls for consistent inconsistency in modelling: modelling principles (e.g. transparency and reproducibility) are standardised, yet the exact method not. If standardisation approaches are implemented on a more regional or local scale the consistent part can be extended, taking into account the consistencies strove for at a higher scale. These consistent parts can change over time based on new insights.

Another potential risk of standardisation is the 'de-skilling' of the experts (Berg, 1997). If standardisation is implemented successfully, the people using that standardisation will hopefully get new skills, while they do not need to use other previous skills. In other words, standardisation users will 're-skill'. However, when there is no focus put on acquiring new skills, it is certain that the standardisation users will end up with less skills than before. Within our case studies, we have not seen this occur. The eWater Source framework has not been implemented completely, meaning that de-skilling is less likely to have happened. Also, training courses have been set up for modellers to obtain new skills using Source. The Dutch guideline is barely actively used, only as a reference, which means that multiple approaches still exist. Even though this potential risk is not visible in our case studies, it is necessary to be aware of it. In the Netherlands, an example that would require re-skilling is the automating of the modelling process as a standardisation approach (**Chapter 3**). The modeller no longer needs to run the model, but needs to run the scripts, resulting potentially in the modeller obtaining less knowledge. While both case studies progress towards the use of a common model framework, the modellers might lose the experience in using other models and thus de-skill.

Lastly, standardisation can be implemented just for psychological benefits (Berg, 1997; Wears, 2015). It provides a rationalist, orderly approach to complex, uncertain systems: '*standardisation is a technological solution to complexity*' (Alder & Wise, 1995; Wears, 2015). For managers and bureaucratic employees, this provides the benefit of the consistency itself. But, for the actual users, there is no apparent advantage to the created consistency from standardisation. In the Australian case, the eWater Source framework contributes to nation-building: it intends to be a nationally consistent framework, which is supported by a community of practice (eWater Source Group, SA). So the framework seems to be partially implemented for social and psychological benefits, while the benefits at modeller level are yet unclear. The Dutch guideline can have a psychological benefit too: if you follow the rules, no one can say you did anything wrong.

All these different risks can result in unintended consequences: modellers might no longer have the same skill level due to deskilling without reskilling, certain methods can become disused because they were not included in the standardisation, and the rewards for the modellers might be minimal even though they put in effort for implementation of a standardisation approach. All of these scenarios mean that standardisation can result in situations where some stakeholders or methods are overlooked or lose value without any apparent reason. Thus, standardisation approaches, like models, are non-neutral tools, requiring careful development, implementation, evolution and reflection upon why it is standardised, who decided upon the standardisation method and who experiences benefits from the standardisation.

4.5 | Conclusion

In this study, we examined the modeller's perspective towards standardisation approaches. We explored this through two case studies: a top-down standardisation approach in Australia and a bottom-up approach in the Netherlands. We conducted 36 interviews: twenty in Australia and sixteen in the Netherlands. The interviewees were modellers using or having used the standardisation approach and developers of the standardisation approach. All interviews were transcribed, after which we executed an inductive content analysis. The analysis resulted in the identification of seven inductive axial codes: Development, Application, Advantages, Disadvantages, Improvements, Risks and Updates. Based on these axial codes, we relayed how the standardisation approaches were perceived in each case study separately. Additionally, we compared the two case studies for these axial codes.

Our results indicate that for standardisation approaches to be implemented, consensus is necessary. Just a decision is not effective to have the approach implemented, though a consensus does not provide a guarantee either. The implementation can be catalysed by a top-down approach or external factors, although site-specificity and different methodologies might impede its development and implementation. The interviewees value the potential benefits of standardisation, such as transparency or reproducibility. Enough resources, proper maintenance and adaptability can promote the necessary evolution of a standardisation approach.

Within this study, we took an individual perspective towards standardisation in the interviews and analysis. However, it is important to include other perspectives, for example as we did through the Human, Organisation, Technology fit framework. Even though standardisation has many advantages, such as transparency and consistency, standardisation also has potential risks and pitfalls. One is the assumption that one approach is the most desirable, reducing the flexibility within the modelling process. To account for site-specificity, modelling can strive for consistent inconsistency, meaning that principles are similar, but the exact method not. The risks and pitfalls of standardisation can lead to situations where certain stakeholders are overlooked or certain methods lose value without any apparent reason. Therefore, standardisation approaches, similar to models, are non-neutral tools, demanding careful development, application, evolution and maintenance.

5

Learning from the critical social sciences

This chapter is based on:

Remmers, J.O.E., ter Horst, R., Nabavi, E., Proske, U., Teuling, A.J., Vos, J., Melsen, L.A. (2024). HESS Opinions: Reflecting and acting on the social aspects of modelling. *Hydrology and Environmental System Sciences (submitted)*

Abstract

Within hydrological modelling, a persistent notion exists that a model is a neutral, objective tool. The notion of neutrality in modelling has several, potentially harmful, consequences. In the critical social sciences, the non-neutrality in methods and research results has been a topic of debate for a longer period already. In order to deal with this in hydrological modelling, we propose that the hydrological modelling network can learn from critical social sciences. This is a call for responsible modelling – modelling that is accountable, transparent and reproducible and this responsibility is carried by all actors related to the modelling study. To support our proposition, we have four pillars of arguments: the social aspects in hydrological modelling, insights from the critical social sciences, building bridges between sciences and reflecting on what the hydrological modelling network can learn. We provide several actionable recommendations as a follow-up. The main take-away, from our perspective, is that responsible modelling is a shared responsibility. We address the complete modelling network – from commissioner, to modeller, to end-user. We invite all actors to take up their share in establishing responsible modelling.

"She never could learn or understand anything
before she was taught;
and sometimes not even then"

—Jane Austen, *Northanger Abbey* (1818)

5.1 | Introduction

Within hydrological modelling, a persistent notion exists that a model is a neutral, objective tool (**Chapter 1,** Frigg & Hartmann, 2024; Savenije, 2009; Wesselink et al., 2017). Although it is generally acknowledged that models influence society, for instance through decision support, this notion of neutrality presumes that the model itself is not influenced by society (Wesselink et al., 2017). Part of this also means that models are deemed to give clear unbiased information for decision support. However, we argue that hydrological modelling takes place within a social context (Krueger et al., 2012; Mayer et al., 2017; Melsen, 2022; Packett et al., 2020), both the problems that are studied with models as well as the modelling process itself (visualised in Fig. 5.1).

The notion of neutrality in modelling has several, potentially harmful, consequences. Neutrality implies that all people and aspects are treated equally. This is not the case (Doorn, 2017; Packett et al., 2020). For example, models are always simplifications of reality, and therefore choices are made on what to represent in the model and what not (Frigg & Hartmann, 2024; Refsgaard, 1996; Savenije, 2009). As a result, the unrepresented processes and aspects are marginalised and become invisible. This can result in injustices: some groups being overlooked, some interest being prioritised, or some ways of understanding sidelined (Doorn, 2017; Zwarteveen & Boelens, 2017).

5

In the critical social sciences – the sciences dealing with critical questions of power relations, especially oppression and domination (Watts & Hodgson, 2019), the non-neutrality in methods and research results has been a topic of debate for a longer period already (Mendelsohn, 1977; Latour, 1990; Law, 2004; Sismondo, 2011). Different disciplines within the critical social sciences, such as Science and Technology Studies (STS) and political ecology, provide different insights into how to deal with non-neutrality. In order to deal with this in modelling, we propose that the hydrological modelling network – all actors, i.e. funders, commissioner, modellers, users, decision-makers, involved in and influencing the modelling study – can learn from critical social sciences. This is a call for responsible modelling – modelling that is accountable, transparent, reproducible and that includes different perspectives – and this responsibility is carried by all actors related to the modelling study.

We are aware that our argument is not new and has been brought up in different terms and ways across the hydrological modelling network. Part of this comes from our own contributions to this debate (**this thesis**, ter Horst et al., 2023; Melsen, 2022; Nabavi, 2022), but we also acknowledge active research communities in Australia working on

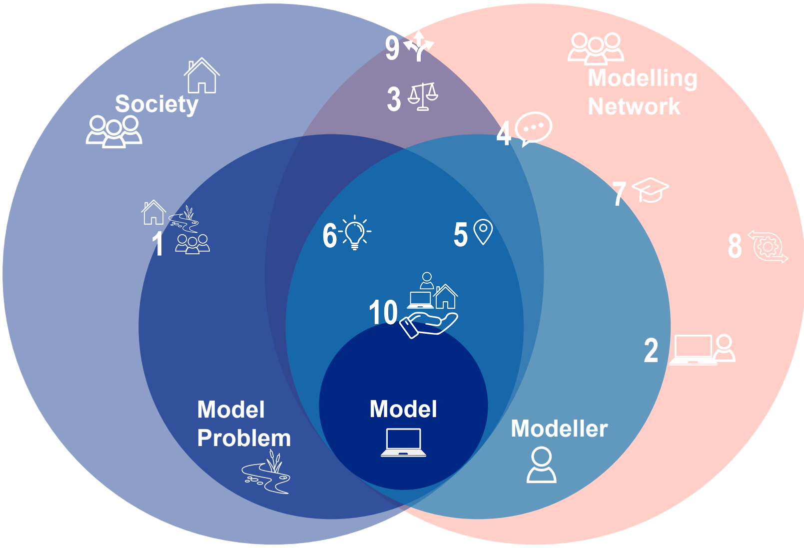

Figure 5.1 | General overview of how the model and its context overlap and where our different arguments are positioned within this. The model problem is the hydrological problem that is being studied and modelled. The modelling network encompasses all actors, i.e. funders, commissioner, modellers, users, decision-makers, involved in and influencing the modelling study. The argument numbers refer to the arguments numbered in the text. 1) hydrological modelling problems are embedded within society, with all its social processes. 2) the modelling process is a social product. 3) the social aspects of hydrological modelling have ethical implications. 4) vocabulary to express the social aspects in hydrological modelling. 5) reflect on positionality and practice active reflexivity. 6) basic understanding of critical social sciences. 7) education can facilitate the knowledge building. 8) structural changes in the modelling network. 9) new avenues for research. 10) take responsibility for model results.

good modelling practices and model governance (Hamilton et al., 2022; Jakeman et al., 2006, 2024), work done in Germany on situated modelling (Klein et al., 2024; Krueger et al., 2012; Krueger & Alba, 2022), ongoing research in France (Molle, 2009; Venot et al., 2014), Post-Normal science (Funtowicz & Ravetz, 1993; Petersen et al., 2011; van der Sluijs, 2002) and sensitivity auditing (Puy et al., 2023; Saltelli & Di Fiore, 2023), work done in the Chesapeake bay (Deitrick et al., 2021; Lim et al., 2023) and the Open Modelling Foundation initiative (OMF, SA). This list is far from complete, but shows the many different aspects that relate to our arguments. That being said, in practice the effects of these studies have been limited (Grimm, 2020), and therefore we provide here a clear overview of arguments to invite the hydrological (modelling) community to start the conversation on the non-neutrality of models.

To support our proposition, we have four pillars of arguments: the social aspects in hydrological modelling, insights from the critical social sciences, building bridges between sciences and reflecting on what the hydrological modelling network can learn. Within each of these pillars, we will provide several subarguments. The main points of the subarguments are highlighted in bold. Figure 5.1 provides an overview of our perspective and where our arguments are positioned in this.

5.2 | Social aspects in hydrological modelling

The first pillar supporting our proposition concerns the social aspects already present in hydrological modelling. Showcasing how hydrological modelling already contains social aspects can highlight the importance for the hydrological modelling network to acknowledge that modelling is not a neutral, purely technical activity. This pillar is underpinned by three arguments.

First, the problems hydrological modellers study are **embedded within society, with all its social processes** (Arg. 1). Hydrological problems relate to society in many ways. Water availability in rivers is impacted by land use changes (Teuling et al., 2019; Wamucii et al., 2021). Unsustainable management of groundwater abstraction has social and political consequences (Nabavi, 2018; Sanz et al., 2019). Or sea level rise necessitates societies to adapt to the risks this brings (Irani et al., 2024; Kopp et al., 2019). These examples led to the initiation of the field of socio-hydrology or hydro-sociology (Sivapalan et al., 2012; Krueger et al., 2016; Melsen et al., 2018b; Ross & Chang, 2020). These disciplines explore hydrological problems as integrated parts of society and often use stakeholder participation as an approach to include the different perspectives to an hydrological problem (ter Horst et al., 2023; Xu et al., 2018). A clear example of the social embeddedness of hydrology is flooding. For instance, the floods that occurred in parts of Germany and the Netherlands in the summer of 2021 caused loss of life (Thieken et al., 2023) and extensive material and economic damages (Kok et al., 2023). This shows how hydrological problems are part of society and can have considerable impacts.

Second, the **modelling process itself is a social product** (Arg. 2), as it inherently contains underdetermined decisions and social processes. Underdetermined decisions arise from equifinality, meaning that certain options are not distinguishable from each other and as such are not 'objectively better' compared to each other (Beven & Freer, 2001; Butts et al., 2004; Ward, 2021; Winsberg, 2012). Although equifinality is often explored in the domain of parameter uncertainty, it can be extended to equifinality in methods or approaches, which might still produce different results or conclusions. Proske et al. (2022, 2023) also use the concept of equifinality when evaluating different

5

model complexities. They show that the simplicity or complexity of the model does not affect the model results substantially – the results are equifinal compared to each other. Equifinality in modelling decisions leaves room for social processes to motivate these, introducing subjectivity and inter-modeller variability (**Chapter 2**, Babel et al., 2019; Krueger et al., 2012; Melsen, 2022). For example, the choice for model software is often based on legacy – the institute a modeller works at determines which software is used (Addor & Melsen, 2019). Additionally, choices made early on in the modelling process can influence choices later on, creating so-called path dependency (Lahtinen et al., 2017; Lenhard & Winsberg, 2010). For example, the chosen model software limits the possible model settings (**Chapter 2**). Furthermore, Lane (2014) argues that the hydrological modeller is not separated from society, and thus is not separated from the problem they study. What these studies show is that the same modelling research question would be answered with a different modelling approach, and therefore likely different model results, at a different time and a different place.

Third, and this is where the previous two arguments come together, **the social aspects of hydrological modelling have ethical implications** (Arg. 3), such as questions about who is involved in the modelling and who benefits (Beck & Krueger, 2016). Due to the modelling decisions and assumptions made (from Arg. 2), model results contain a specific perspective of reality (Nabavi, 2022; Saltelli & Di Fiore, 2023). Choosing this perspective means excluding or sidelining other perspectives. Stakeholder engagement has the potential to bring these marginalised perspectives forward again (Packett et al., 2020; Xu et al., 2018), although stakeholder engagement comes again with its own challenges (e.g. Reed et al., 2009; Turnhout et al., 2020). As model results have societal implications, injustices can occur (Thaler, 2021; Zwarteveen & Boelens, 2017). For example, models used in flood studies, which obviously can have high societal impact, can cause injustices. They might for instance not consider informal settlements in the floodplains, and as such marginalise inhabitants of those floodplains (Wesselink et al., 2017).

5.3 | Insights from critical social sciences

Critical social sciences provide the tools and theoretical frameworks that can address the social aspects of hydrological modelling. Here, we will highlight three.

First, the critical social sciences have the **vocabulary to express the social aspects in hydrological modelling**. Different disciplines of critical social sciences can provide various suggestions. This vocabulary is not (yet) common in the hydrological modelling network, even though similar concepts are addressed, albeit described more elaborately. For example, when we just described '*model results contain a specific perspective of reality*' in the previous section (in Arg. 3), we could have also used the term 'situated', which is also used in feminist theories (Haraway, 2013). This means that model results are formed in a specific context. Another example is 'ontology', meaning the study of the existence of things and what they look like (Frigg & Hartmann, 2024; Wesselink et al., 2017). With a model, a researcher studies what a hydrological system looks like. The representations a researcher chooses are dependent on their ontological view of the system. In more recent literature (e.g. Klein et al., 2024; Moon & Blackman, 2014; Wesselink et al., 2017), some of the critical social science vocabulary and concepts are related to the environmental sciences. Knowledge of this vocabulary can enhance our understanding of and facilitate our discussion of the social aspects in hydrological modelling (Laplane et al., 2019).

Second, social scientists **often reflect on their positionality and practice active reflexivity** in their research (Arg. 5). A positionality statement is written to indicate how they as researcher relate to the subject they study (Lin, 2015; Njeri, 2021; Soedirgo & Glas, 2020). For example, critical social science disciplines using ethnographical methods – observing subjects in their own environment – often include a positionality statement, since the scientist's background influences the observations and interpretations they make. Hydrological modellers also have a personal perspective/position (from Arg. 2) towards their subject through their own previous experience or the institute they work at or even their own personal interests and hobbies (Deitrick et al., 2021; Melsen, 2022; Packett et al., 2020). Based on these experiences or contextual factors, modellers tend to make decisions (**Chapter 2**, Krueger et al., 2012; Melsen, 2022; Sanz et al., 2019). Reflecting on and being transparent about positionality can create more transparency regarding this personal context and assumptions made (Blackett et al., 2024; Klein et al., 2024; Wesselink et al., 2017). For example, Melsen (2022) includes a brief positionality statement for the interview study she did, highlighting how her own background has influenced the conducted interviews. Besides writing a positionality statement, active reflexivity – continual questioning of your own assumptions and biases – should also be done throughout the modelling process (Soedirgo & Glas, 2020). This entails documenting assumptions, normalising reflexivity, engaging others in the reflexivity and publishing reflexivity alongside the research. Publishing reflexivity through a positionality statement means being vulnerable and open. This vulnerability can build trust in how models are used, because it is more transparent. Additionally, it can inspire others to also reflect on or to become more open about their modelling practices and assumptions. As more people start to do this, it can slowly change practices in the whole modelling network.

Third, again combining the previous two arguments, **basic understanding of critical social sciences** is needed to situate your own research in a broader context to understand the possible positive and negative consequences of modelling and to be able to identify who to empower and how (Arg. 6). This context is needed since hydrological modelling addresses societal issues (from Arg. 1), the hydrological modelling process is a social product (from Arg. 2) and the ethical implications model results have (from Arg. 3). The necessary basic knowledge should entail knowledge to place modelling results in the societal context (from Arg. 1) and reflect on potential ethical consequences of the results (from Arg. 3), for example knowledge on flood warning responses to understand what model results mean. In 1997, the National Weather Service did not include the uncertainties when they issued a flood warning two months in advance for the Red River, North Dakota, USA. Because of this, a town, Grand Forks, thought it was safe. But, the actual flood reached the upper band of the uncertainty range, resulting in flooding of the town and 75% of the houses were damaged. Currently, the National Weather Service does provide that information (Silver, 2012). Thus, executing an uncertainty analysis or not as a modeller can have ethical implications in society in water management (McMillan et al., 2017; Silver, 2012). Recently, ethics of Artificial Intelligence has gained traction (Doorn, 2021; Maier et al., 2024; Nabavi et al., 2024) and rightly so. However, surprisingly, very little is done on the ethics of numerical (hydrological) modelling. Understanding of certain concepts of critical social sciences can also ease reflecting on the subjectivity in modelling (form Arg. 2). For instance, the vocabulary (from Arg. 4) can help expressing the subjectivity or help initiating reflexivity. Ontology – studying what something looks like – can spark debate on the different perspective people have of a hydrological system (Verzijl et al., 2023). A person living somewhere can define what a system looks like differently than a researcher or tourist: more as a social, physical or technical thing.

5.4 | Building bridges between sciences

Different researchers have been trying to build bridges between disciplines (Krueger et al., 2016; Pulkkinen et al., 2022; Ross & Chang, 2020; Rödder et al., 2020; Venot et al., 2022; Zwarteveen & Boelens, 2017), but most remain within their own discipline. Hierarchy of sciences – the idea that certain sciences, such as physics, have a higher degree of consensus and scientific advancement than others, such as social sciences – reinforces this way of thinking and acting (Comte, 1855; Cole, 1983; Fanelli, 2010; Simonton, 2006). We propose two ways in which the hydrological modelling network can increase the building of bridges to critical social sciences: first, through education, which will instigate structural changes in the long-term and, second, through structural changes that can have an immediate effect.

First, **education can facilitate the knowledge building** necessary to understand the basic critical social science concepts (Arg. 7). Understanding basics of other sciences can increase communication and effectiveness in future work situations, enhancing inter-disciplinary collaborations (from Arg. 6). This knowledge building of social processes and reflexivity needs to be practical and integrated within hydrological modelling education (Micheletti et al., 2024; Oldfield, 2022; Stefanidou & Skordoulis, 2014). For example, the curriculum for hydrological modelling education should have reflexivity and responsible modelling integrated in its curriculum: within a modelling course, the students learn to apply reflexivity as they model. Education should extend to working professionals in order to keep up with new insights and to also incorporate this knowledge in the current workforce.

Second, although education can help raise a new generation of hydrological modellers, we need **structural changes in the modelling network** to facilitate the incorporation of social aspects in daily modelling practices (Arg. 8). Structural changes can guide and force the hydrological modelling network to adapt practices focusing on taking the social aspects into account (Jakeman et al., 2024). For example, funding requirements can include a positionality statement within the funding application (from Arg. 5) or a research plan that specifically designates time for active reflexivity. Also, journal requirements can be adapted to incorporate social aspects in hydrological modelling more explicitly. Journals might start asking for a positionality statement as well, or they can ask for documentation on assumptions in the modelling process.

5.5 | Reflecting on what the hydrological modelling network can learn

Building a bridge to critical social sciences can improve transparency about the social aspects of hydrological modelling. Also, considering and disclosing the uncertainties associated with these aspects potentially creates more reproducibility. Increased transparency and reproducibility can contribute to more productive scientific progress and responsible policy making that can be justified.

Also, acknowledging social aspects in hydrological modelling can open **new avenues for research** (Arg. 9). Critical social science understanding can move the hydrological modelling network towards more productively working on societal problems (from Arg. 7). Through reflecting, modellers are incentivised to rethink their modelling decisions. This might result in more robust and accountable modelling decisions. In turn, this will provide more accountable decision support. Reflexivity highlights assumptions made. Sharing these assumptions can streamline research with less redoing of each other's work (Laplane et al., 2019). It is easier to know what has or has not been done before and to have the ability to complement each other because of that knowledge. Different researchers would facilitate diversity in approaches and therefore give a more complete picture (Baldissera Pacchetti et al., 2024). Additionally, it could be that new research will specifically look for diversity, instead of a universal model (Baldissera Pacchetti et al., 2024; Horton et al., 2022; Savenije, 2009). Flexibility can also be introduced through modular modelling frameworks (Clark et al., 2008; Craig et al., 2020; Fenicia et al., 2011). This diversity can encompass the different contexts in which the modelling is shaped or in which the modelling is used.

With more transparency on the social aspects of hydrological modelling, modellers and also funders, commissioners and decision makers can **take responsibility for model results** (Arg. 10). This should be a shared responsibility, not just the modeller's. The interplay between these actors can create dynamics influencing the modelling. This interplay should be made more visible (from Arg. 5). Structural changes in the modelling network (from Arg. 8) can facilitate this. Due to the transparency, modelling results will be more retraceable, and the limitations of a modelling study are more evident. Including the social context within the modelling process, can provide better information for decision/policy makers. Having modellers reflect on their ontology of a hydrological system will help in their ability to recognise how their model results are partial and might have looked different with another ontology. Being aware and transparent of this contributes to decision makers being able to justify their policy decisions. For instance, after flooding in Brisbane and surrounding, the model results were questioned and the organisations behind them were held responsible (Supreme Court of New South Wales, 2021). This example shows that the organisations using and providing model results need to be able to take responsibility of them. Sharing responsibilities can take many forms, but it starts with curiosity for and openness to knowing, understanding and taking action on the social aspects of hydrological modelling. Another example, outside of hydrology, is that the modellers that simulated the nitrogen emissions for a newly planned airport in the Netherlands were investigated by the Public Prosecution Service, because there were clear indications that all modelling decisions were made such that the nitrogen emission was as low as possible (Adecs Airinfra Consultants, 2021; NOS Nieuws, 2022). Not surprising perhaps, if the executing company sells themselves as "aviation lovers", but also the result of a commissioner that has certain interests. As such it is a clear example of how modellers can be held accountable for their model results, while they also face forces from, for instance, funders.

5.6 | Invitation to start acting

From the arguments we have provided in the previous sections, we can derive possible tangible actions that can be implemented by different actors in the modelling network. This list of actions is not exhaustive. As potential follow-up actions, we recommend:

- If you are a model user (i.e. someone who analyses and uses model results), you can consider asking the modeller for their assumptions and how the modelling process was executed (based on Arg. 2).
- If you are a modeller, you can consider to start reflecting on your positionality and consider to include a positionality statement in your next modelling study. How did your experience and position in society influence how you approached this study? (based on Arg. 5)
- If you are teaching the next generation of hydrological modellers, you can consider incorporating reflexivity practices or critical social science vocabulary and concepts in your lecture, computer practical, course, or curriculum (based on Arg. 4, 5, 6 and 7). For instance, York et al. (SA) are working on a so-called 'recipe' book, from which teachers can drawn inspiration for incorporating STS (one of the critical social science disciplines) modules in their course. We, as the authors, were not able to find many or any examples in actual courses. Therefore, we would also recommend sharing different experiences with or examples of incorporating this in courses.
- If you are a commissioner, you can consider allowing for more time or funding during projects for including reflexivity in the modelling process or writing a positionality statement (based on Arg. 5 and 8). You can also consider to change your project requirements to include reflecting on positionality.
- If you are overseeing a modelling team, you can consider having a discussion on internalised assumptions in your way of working, also known as entrenched workflows (based on Arg. 2, Levine & Wilson, 2013), and on the potential implications of your modelling work (based on Arg. 3).

These follow-up actions sound like a recipe. However, in this whole opinion paper, we have advocated and shown that hydrological modelling is context dependent. Therefore, we acknowledge that anyone implementing these potential actions needs to navigate their own working environment. More importantly, this list is not definitive; we invite you to explore and discuss this topic further and come up with your own ways of incorporating the different insights from the critical social sciences.

5.7 | Conclusion

In this chapter, we argue why and how we think the hydrological modelling network, which we define as all actors, i.e. funders, commissioner, modellers, users, decision-makers, involved in and influencing the modelling study, can benefit from some practices used in the critical social sciences. To support this, we have four pillars of arguments: the social aspects in hydrological modelling, insights from critical social sciences, building bridges between sciences and reflecting on what the hydrological modelling network can learn. Based on these arguments, we provide some tangible follow-up actions targeting the whole modelling network to promote responsible modelling – modelling that is accountable, transparent and reproducible. This responsibility is carried by all actors related to the modelling study. Even though we focused on the hydrological modelling network, these lessons are also applicable to other modelling communities.

The main take-away, from our perspective, is that responsible modelling is a shared responsibility. We realise that modellers tend to already bear a lot of the responsibility and are the easiest ones to ask actions from. Substantial change is not possible without also addressing the other actors in modelling studies, such as educators, funders or supervisors. Therefore, we address the complete modelling network. We invite all actors to take up their share in establishing responsible modelling.

6

Synthesis

"Every qualification is raised at times,
by the circumstances of the moment,
to more than its real value."

—Jane Austen, *Sense and Sensibility* (1811)

6.1 | Summarising this thesis

In this thesis, I explored how hydrodynamic decision-support modellers motivate modelling decisions and how they perceive standardisation approaches of the modelling process. Inter-modeller variability introduces differences in model results. To be able to put the results in perspective, it is important to understand the modeller's motivations behind their modelling decisions. Standardisation approaches are being implemented to counter this variability. Since modellers need to implement these approaches in their practice, the question is how they perceive them.

First, I analysed the motivations behind modelling decisions in **Chapter 2**. Second, I studied how modellers perceive automation as a standardisation approach (**Chapter 3**). Third, I researched how modellers perceive bottom-up and top-down standardisation approaches (**Chapter 4**), respectively a Dutch modelling guideline and an Australian common modelling framework. In **Chapter 5**, I combined my findings and insights of **Chapters 2 - 4** to give recommendations for moving forward in the hydrological modelling community.

Because the water management context will vary, so will the considerations and motivations. Deitrick et al. (2021), also, indicated that other case studies will improve our general understanding of this topic. I have explored one case study with regards to modelling practices and two concerning modelling standards. Furthermore, I have interpreted and analysed these case studies from my own perspective and with my own background and experience, but other researchers might have done this differently. Hence, it might be good to explore other water management contexts by other researchers. Additionally, I only analysed the modelling practices at the time of the interviews, which can also be extended with exploring how motivations can change over a modeller's career. Regarding standardisation approaches, I explored the modeller's perspective of this at one point in time. However, standardisation approaches will likely evolve over time in response to modellers' perceptions. Thus, it would be interesting to analyse how a modeller's perspective and a standardisation approach change and evolve together.

Finally, in this chapter, I synthesise and combine these different findings. First, I relate both the modelling practices and standardisation to each other. Second, I reflect on positionality, both my own during this PhD and postitionality generally in modelling. Third, I will look at how models are used in decision support. Finally, I will conclude this thesis by evaluating 'Modellers as Influencers'.

6.2 | Combining practices and standards

Within this thesis, I have mostly addressed my findings on practices (**Chapter 2**) and standards (**Chapter 3** and 4) separately. I have built my reasoning for looking into standardisation on the fact that inter-modeller variability exists in modelling practices and increasing the justifiability and transparency of modelling results. This connection from inter-modeller variability to standardisation for more justifiable and transparent model results indicates a link between practices and standards. Moreover, as I have argued before, I analysed how modellers perceive standardisation, because it influences their practices. As such, it is obvious multiple links exist between practices and standards. Therefore, my findings also need to be related.

First, I found that social aspects play a role in practices and standards (mainly subjectivism). **Chapter 2** shows how the individual modeller and the modelling team are the main motivation categories for modelling decisions, influencing the modelling practices. The implementation of standardisation approaches depends partially on how the modellers implement and perceive it (**Chapters 3 and 4**). As much as models can marginalise certain perspectives on the world by simplifying certain aspects of the real world, standardisation can marginalise certain modelling practices by making a certain practice the desired one.

Second, my findings show that the broader context proves to be important in practices and standards (relativism). For the practices, the individual and team motivation behind modelling decisions are limited by other considerations in the modeller's broader context, such as organisations or national agreements (**Chapter 2**). For the standardisation approaches, their development and implementation depends partially on external factors, such as resources for maintenance or enforcement by law, because these can catalyse these (**Chapter 4**). Modelling activities are not solely determined by modellers and their model, but are part of a broader context that shapes the boundary conditions.

Third, both practices and standards show forms of consistent inconsistency. An individual modeller often relies on their own common practices and experiences to apply consistency in their modelling. But, they do take the differences for each modelling project into account (inconsistency), adapting their modelling strategies to it (**Chapter 2**). In **Chapter 4**, I argue that standardisation should take site-specificity into account (inconsistency), but a similar modelling practice is striven for through e.g. standardisation (consistency). More transparent and justifiable modelling is partially balancing the consistency and inconsistency within each modelling study.

Finally, modelling and standardisation are both forms of pattern seeking (mainly constructivism). Asking for consistency, for example through standardisation, is asking for a pattern and allow the developers and users of standardisation approaches to remain in the technical aspects of modelling (Alder & Wise, 1995; Wears, 2015). In modelling, a pattern is sought through the representation of the hydrological system chosen. Pattern seeking is a common characteristic within modelling activities.

I think it is important to recognise that human-beings seek patterns in everything, because it makes the complex, uncertain world more comprehensible (Andrienko et al., 2022; Kim et al., 2020; Leamer, 2009). I do this as well within this thesis: in every content chapter (**Chapters 2 - 5**) I have made a graphic or illustration, which is organising my findings as comprehensibly as possible in a pattern I perceived. Recognising patterns is done from an individual's perspective on objects, data or processes. Savenije (2009) and Beven (2007) pointed out that hydrologists cannot observe the hydrological process at the same scale as our model represent them. To be able to observe them, we need to zoom out. Zooming out is also required to recognise our own patterns, assumptions and biases in our modelling practices. This can improve understanding of and potentially lead to changes in our patterns, assumptions and biases. Hence, unless the broader context and the personal influence are acknowledged in practices and standards (zooming out), hydrological modelling will stay in a positivist mindset.

6.3 | Reflecting on positionality

6.3.1 | My own positionality

To zoom out for my own thesis, I will reflect here on my own personal influences and context. These have influenced how I executed the research for this thesis.

Being Dutch and growing up in the Netherlands resulted in me knowing the Dutch system fairly well (subjectivism and relativism). This aided me in preparing and conducting the interviewees in the Netherlands. But for the Australian interviewees, I had to put in extra effort to understand their system. This insider or outsider perspective has influenced the questions I was able to ask and the level of detail I was able to understand and question about. During the Dutch interviews, I was able to ask more specific questions more quickly, which meant it was sufficient to conduct less interviews. However, I might have sometimes not asked the follow-up question, because I made the assumption of already understanding (partially also to keep time). I think it is possible to do this research both as an outsider and insider, because both perspectives bring their own opportunities and challenges. As an insider, I have to keep questioning my assumptions, but it is easier to connect with interviewees and prepare for the interviews. As an outsider, I do have to put in more effort to understand the system and align the interviews. For example, for my Australian interviews, I took a road trip with different guided tours by local experts to understand the hydrological system better. Also, I had some informal meetings with experts to improve my understanding of the physical and political system. Finally, my native language is Dutch, but my education has been mostly in English. Therefore, in some ways, the Dutch interview were slightly easier with just general conversation, but the Australian interviews were easier regarding the hydrological jargon.

Throughout my studies, I focused on climate and water, studying both the environmental and social aspects (subjectivism). In my BSc thesis, I conducted several interviews, while in my MSc thesis, I executed a modelling study (Remmers et al., 2020). Immediately after my MSc, I started wit my PhD at the Hydrology and Environmental Hydraulics group at Wageningen University. Having this mixed background allowed me to 1) connect with the interviewees on what it means to use models and 2) build on my basic knowledge of conducting interviews. However, this also meant that I am not a complete expert on both subjects. Still, in some interviews, this gave me the opportunity to let the interviewee explain more and use my (assumed) lack of knowledge to my advantage.

Due to my questions, some of the interviewees mentioned at the end of the interview that it was informative, for example, regarding the reflection on their own modelling decisions or in the casual talk, not included in the interview, to hear about my work. Most of the interviewees also have a natural science background with a positivist viewpoint. Hence, reflecting during the interview might have changed their perspective slightly. But, I do not think I interviewed the strict positivist modellers, because those probably would not have wanted to participate in my interviews. One way to reach potential interviewees was to pose an open question in the hydrodynamic modelling project I was involved in. Besides this, I targeted people personally who I, my supervisor and co-authors thought would be willing to participate. Still, one person I personally emailed did refuse to participate, because they did not think my research would have any added value.

Not being a full expert in one of the social sciences, I do not have the necessary (detailed) knowledge of the social sciences and the relevant concepts for my work (subjectivism). As such, I am a great example for my argument in **Chapter 5**: I would have benefited from having known more concepts and practices of the critical social sciences. Unfortunately for me, no one in my direct working environment had this expertise either. I had to actively search for and learn these concepts. Even at the time of writing this, I have not completely taken up this necessary knowledge; I would still need a cheat sheet. For example, I enlisted the help of a science philosophy professor to check Section 1.3 on its correctness. For **Chapter 5**, I have enjoyed bringing together the different perspectives of all my co-authors, since all of them have a certain expertise that I do not have. The social scientists as co-authors complemented my expertise very well in that regard. For me, it is still an ongoing process to take up the insights I detailed in **Chapter 5**, such as learning how to actively reflect, write a positionality statement or learning the basic concepts.

Finally, working in a hydrological group and studying physical sciences (relativism) biased my way of thinking at the beginning towards hydrological modelling and the related assumptions (positivist, post-positivist mindset and results are objective/neutral), as I explained in **Chapter 1**. However, due to my experiences throughout and research for my PhD (subjectivism and constructivism), I have moved away from the (post-)positivist perspective. For example, by interacting with other experts within and outside my own discipline, I have seen how research is shaped by the individual and by the individual's environment, and how research is constructed. Even just conducting the interviews has changed my mindset to be more questioning about how I do things and make my decisions. This can have a paralysing effect (not recommendable). Therefore, I think a healthy dose of pragmatism is also necessary, but I have not mastered this quite yet either. Currently, I find myself moving between three philosophies: Constructivism, Subjectivism and Relativism with still a hint of post-positivism.

6.3.2 | Positionality in modelling

As argued in **Chapter 5**, modellers can also write a (short) positionality accompanying their modelling study. A positionality can elucidate their own stance towards a project (Lin, 2015; Njeri, 2021; Soedirgo & Glas, 2020). I have found it a valuable tool to reflect on my own stance and assumptions, though I think it would have been useful to have been more actively reflexive throughout my PhD. For example, before doing my research stay in Australia, I think it would have been helpful to more explicitly state my outsider perspective for myself. I was not familiar with this way of working, and I did not put in the effort to learn and apply it during my PhD. Even now, as I am writing this positionality and reflecting on my background and positionality, I have mainly touched upon how I as an individual influenced my PhD (subjectivism and constructivism).

There are also processes in the broader context that shape everyone's philosophical world views or in this thesis, modelling point of views (relativism). One way to think about the broader context in a positionality is as '*sedimentation*' (Merleau-Ponty, 1945). Similar to a riverbed forming sedimentary layers because of the sediment that moves through, values and beliefs become 'sedimented' in (modelling) cultures. A new value or belief forms a new sediment layer in the culture, because they are introduced by new sediment at one point. Over time, multiple values or beliefs become embedded in the culture or riverbed on top of and mixed through each other. The people growing up in that culture – in this case, the people being educated or working in that culture – internalise these values and beliefs in their own thoughts and ways of working. Another way to describe this is workflow entrenchment (Levine & Wilson, 2013). In hydrology, one example is the existence of different schools in how detailed hydrological models ideally should be: ranging from bucket models (Perrin et al., 2003) to 3D grid models (Samaniego et al., 2010a,b). Other examples of entrenchment are that calibration is a standard procedure or that streamflow is the main focus of many hydrological studies. Being transparent about your modelling school can be a part of a positionality. For the people in these schools and cultures, it might seem that the riverbed is unchangeable. However, exchanging values and beliefs – each others sedimentary history – will expose the individual's assumptions in research and modelling studies.

Being transparent about your own assumptions and personal background can lead to the idea that anything would go, in the sense that being transparent will give a pass for whatever is done. This clash between acknowledging the situation and anything goes is shown in the debate around situation ethics (Dupre, 1967; Munteanu et al., 2015; Raskin & Debany, 2018). Proponents claimed that the ethical thing to do in any situation is completely dependent on the specific situation – taking into account the characteristics and needs of that specific situation. Translated to modelling, this would mean that a modeller can tailor their modelling process based on the requirements of the modelling problem. Opponents of situation ethics claimed that this type of ethics or reasoning could be used to justify that '*anything goes*'. In modelling, *anything goes* can mean that any model, tool or script can be used in any situation. However, this is not how responsible modelling is executed, because modelling decisions need to make sense within the given situation – they need to be fit-for-purpose (Hamilton et al., 2022). The fit-for-purpose framework acknowledges that modelling is situational, but it does say it needs to fit the local purpose. As such, it counters the notion that '*anything goes*' in hydrological modelling. The interviewees in **Chapter 2** also highlighted this by motivating their modelling decision with '*Depends on the area/goal/project/situation*'. This indicates that for them modelling decisions need to fit the purpose of the specific modelling study. However, making these decisions to create situation-dependent fit-for-purpose models does not take into account inter-modeller variability, so in the same situation, modellers still make different decisions. Writing a positionality statement is one tool to improve transparency in the context of and the accountability of the model results.

However, most modellers, like I was, are not familiar with positionality statements or even reflexive modelling (Glattfelder, 2019; Klein et al., 2024; Laplane et al., 2019). This does reiterate the point of integrating these topics in the curriculum of hydrological modelling (**Chapter 5**). Reflexive modelling can also accentuate during the modelling process what choices modellers make. This might prompt modellers to reexamine their decisions, which might lead to more accountable and transparent modelling used for decision support.

6.4 | Using decision-support models

In water management, hydrological model results are used to support decision making. As models have been used frequently and for a long period (Burt & McDonnell, 2015; Horton et al., 2022), a model's usefulness can be taken for granted and its usability taken as it is (Devi et al., 2015). However, in applying a decision-support model now, using it is the foremost driver (Savenije, 2009). Useful, usable and used are three terms that are often thought of in succession: first useful, then usable and finally used (Aitsi-Selmi et al., 2016; Boaz & Hayden, 2002).

In decision-support modelling, however, useful, usable and used are not explicitly called for when applying models or certain modelling tools. The fact that models are so frequently applied (Burt & McDonnell, 2015) strengthens general perceived usefulness of them, and consequently increases their even more frequent, and perhaps unquestioned, usage further. Of course, there still needs to be some usefulness and usability present in the beginning, but it mainly starts with the fact that the modelling is deemed to be required and to be used. For example, **Chapter 2** shows that feasibility, for instance the consideration of time and availability, was a more important consideration than reliability or usefulness. This focus on feasibility shows that models usage is an important driver in making modelling decisions. External factors, such as legacy (Addor & Melsen, 2019) or data availability (Melsen, 2022), limit the capacity of prioritising useful and usable before used. Nielsen (1995) studied how usability of tools translates to them being used. He found that costs of the tool and having a specific need for a tool were determining factors in whether a tool was used or not. So, useful and usable are not a the forefront initially when executing a modelling study for decision support.

After or as a modelling tool is adopted in modelling practices, it can be developed further to increase its usefulness and usability. **Chapter 3** highlights that automation is often developed more crudely in the beginning to use right in the moment. Afterwards, scripts are polished up and made more usable, as their usefulness is proven further. In the Australian case study of **Chapter 4**, the initial users of the common modelling framework had to report bugs and errors to improve the usability. Additionally, some features were initially not present in the framework, which made the framework seem less useful for some users. As these features were implemented, the usefulness increased, but the general usability decreased, as the framework became computationally heavy to run. So, the framework is used due to the top-down approach, but the usefulness and usability are still being improved on. This also demonstrates that the perspective on usefulness and usability of models differs between individuals and organisations.

However, perhaps different levels of usefulness and usability are enough for creating accountability in decision-support modelling depending on the situation. Savenije (2009) noted that modellers in water management have a different purpose associated with modelling than scientific modellers. Scientific modellers want to understand what is wrong with a model and want to improve it. While in water management, models are viewed as pragmatic tools, which need to have been established as reliable in the decision-support modelling network. For decision support, accountability and justifiability of the modelling results also play an important role, which usefulness and usability can impact. Usefulness can increase the justifiability through showing better how a model can be fit-for-purpose. Usability can increase accountability through potentially creating a more transparent and explainable modelling process. However, if models are perceived as pragmatic tools in decision-support modelling (Savenije, 2009), then maybe pragmatism is another factor to be considered besides usefulness and usability in decision-support modelling.

Standardisation can also be implemented to increase justifiability and accountability, as shown in **Chapters 3 and 4**. In theory, I am not against implementing standardisation as it potentially creates transparency, reproducibility and consistency (**Chapters 1, 3 and 4**). These aspects can all aid in improving justifiability and accountability of modelling results. But, I am sceptical how it will work out in practice, especially in the long term, because **Chapters 3 and 4** show that certain requirements need to be met, such as good documentation and ample resources. Furthermore, standardisation is also a social construct and has social implications through for example marginalisation of both modelling practices and consequently perspectives of the real world through simplifications in the 'desired' modelling practice. Therefore, standardisation is a non-neutral tool (**Chapters 3 and 4**) and it is not a panacea. As such, I would recommend to be careful with its development and implementation.

As I have reflected on useful, usable and used models, I like to do the same for this thesis. It is useful: it has highlighted the social aspects in modelling practice and standards. Its usability is probably quite low. In **Chapter 5**, I give some concrete suggestions for how to use critical social sciences in the hydrological modelling community. This (partially) covers how to deal with social aspects in modelling practices, but not in standards. Currently, my findings are not used. I could be kind and say '*Give it more time*', because I have only just finished this PhD thesis. However, to be realistic, creating science that is used, especially science challenging the status quo, is quite challenging.

So, after reflecting on useful, usable and used aspects of my thesis, I would recommend increasing the usability and implementation of my thesis. This would require coming up with actionable points beside those in **Chapter 5** and spreading the word on these usefulness and usability of these action points. Including other experts besides modellers and social scientists, such as outreach coordinators, implementation strategists and organisational leaders, might help tackle the different challenges along the way. This would call for a lot of funding and time, because the people involved need to get some financial compensation and organising all these different experts and making things usable and used will require an iterative process of creating, implementing and reflecting on possible action points. Therefore, I do not think it is very realistic, but perhaps if it is addressed one step at time. And maybe, due to my interactions with others during my PhD, I might have planted some seeds for them to incorporate small parts in their daily practice.

6.5 | Modellers as influencers?

But even with models established as tools in decision support, modellers can act as influencers – the ones influencing. So, how can modellers act as influencers? I will explore this question from three points of view: modellers as infuencers within the modelling process, on decision making and when zooming out to society.

First, just looking at the modeller and the model combined, the modeller can act as an influencer. As **Chapter 2** shows the Individual motivation category is the largest of the eight motivation categories. Henckens & Engel (2017) showed in a benchmark study that the influence of modellers is larger than the influence of the model. Still, it is not the only motivation category and more importantly, their sphere of influence in the modelling process is largest in the model implementation step. In other modelling steps, such as choosing model software or forcing data, the other motivation categories are more influential. Also, for standardisation, the modeller influences how they implement it and they have some influence on what the standardisation approach might look like (**Chapter 3 and 4**). However, **Chapter 3** shows that automation can only be seen as a standardisation approach at the (inter-)organisational level. Additionally, **Chapter 4** shows that external factors, such as governmental pressures, are necessary to develop and implement any standardisation approach. These chapters clearly show that, indeed, modellers can act as influencers within their own sphere, but this is limited by external factors and considerations.

Moving towards a broader perspective, where model results are used for decision-support, modellers do retain some influence. Model results are considered by decision makers. On top of that, modellers determine how their model results are interpreted and presented. However, many other aspects besides hydrological model results are also considered in decision making, such as time and resources available, other model results from different disciplines and different stakeholder priorities (Darling, 1945; Connell & Grafton, 2011). Thus, modellers still have a sphere of influence on decision making through the model results that support this. However, this is substantially limited by other societal factors and priorities.

From a more negative point of view, it can even be said that modellers are influenced. For example, the philosophical concept sedimentation explains how one grain in a river is influenced by the river itself, such as its history or the other grains, while at the same time also influencing the course of the river. In a similar fashion, a modeller is influenced by the team and organisation they work in – the other grains – through internalising the common practices. Also, a modeller is influenced through their own experiences – their history, such as their education and previous working situations. And finally, a modeller is influenced by society itself - where modelling is a generally accepted method, funding and power are redistributed, and certain ideas or concepts are prioritised. The Lelystad airport example, described in **Section 5.5**, clearly shows that certain interests of the people who fund the study are served. Due to these influences, a modeller's modelling practices are shaped. Therefore, one could even say that modellers are influencees – the ones being influenced, since society and the modelling ecosystem shape the modellers maybe even more than the modeller shapes the model. As such, in summary, modellers have their own sphere of influence in the modelling process and outside of it, but this is bound and influenced by the societal context.

7

Supplementary Information

Appendices

General appendix: A - Interview guides

Interview guide - Modellers NL

The interview guide consisted of three parts: an introductory part covering the background of the interviewee, a specific part covering the full modelling process and a final part concerning general question about modelling. The questions are subdivided in main questions and subquestions to probe the interviewee if necessary.

Questions regarding the background:

1. How would you describe your current position?
2. What is your background?
   a) Have you worked at other companies/institutes before?
   b) Where and what did you study?
3. Can you describe your experience as a modeller?
   a) Which type of models?
   b) Which models models exactly?
   c) How many years?
4. For which type of questions do you use a model?

For interviewees from water authorities, we also asked how much and which parts of the modelling were done in-house or externally.

Questions about the full modelling process:

Model Software

1. Which software do you generally use?

2. How did you make this decision?
    a) Why this software?
    b) Are there other options?
    c) Why for this purpose?
3. Did you have certain settings in your model?
    a) If no:
        i. Why not?
        ii. When would you have certain settings?
        iii. If this step were to be automated would you use it / execute this step?
            A. Would you find this useful?
            B. Why (not)?
    b) If yes:
        i. Why these settings?
        ii. In which situation(s) might you chose different settings?
        iii. Do you think this step can be automated in the future?
            A. Why (not)?
            B. To what extent?
            C. What are possibilities for automation?
4. In which situation(s) might you chose a different software?

Forcing Data

1. Which forcing data do you generally use?
    a) Type?
    b) Source?
    c) Resolution of data, spatial and temporal?
    d) Availability?
2. How did you make this decision?
3. In which situation(s) might you chose different data?

4. Do you think this step can be automated in the future?
    a) Why (not)?
    b) To what extent?
    c) What are possibilities for automation?
5. Do you perform any pre-processing on the forcing data?
    a) If no:
        i. Why not?
        ii. When would you pre-process the data?
        iii. If this step were to be automated would you use it / execute this step?
            A. Would you find this useful?
            B. Why (not)?
    b) If yes:
        i. How do you generally execute any pre-processing?
        ii. How did you make this decision?
        iii. In which situation(s) might you chose a different method?
        iv. Do you think this step can be automated in the future?
            A. Why (not)?
            B. To what extent?
            C. What are possibilities for automation?

Static Data

1. Which static data do you generally use?
    a) Type?
    b) Source?
    c) Resolution of data, spatial and temporal?
    d) Availability?
2. How did you make this decision?
3. In which situation(s) might you chose different data?

4. Do you think this step can be automated in the future?
    a) Why (not)?
    b) To what extent?
    c) What are possibilities for automation?
5. Do you perform any pre-processing on the static data?
    a) If no:
        i. Why not?
        ii. When would you pre-process the data?
        iii. If this step were to be automated would you use it / execute this step?
            A. Would you find this useful?
            B. Why (not)?
    b) If yes:
        i. How do you generally execute any pre-processing?
        ii. How did you make this decision?
        iii. In which situation(s) might you chose a different method?
        iv. Do you think this step can be automated in the future?
            A. Why (not)?
            B. To what extent?
            C. What are possibilities for automation?

Observation Data

1. Which observation data do you generally use?
    a) Type?
    b) Source?
    c) Resolution of data, spatial and temporal?
    d) Availability?
2. How did you make this decision?
3. In which situation(s) might you chose different data

4. Do you think this step can be automated in the future?
   a) Why (not)?
   b) To what extent?
   c) What are possibilities for automation?
5. Do you perform any pre-processing on the observation data?
   a) If no:
      i. Why not?
      ii. When would you pre-process the data?
      iii. If this step were to be automated would you use it / execute this step?
         A. Would you find this useful?
         B. Why (not)?
   b) If yes:
      i. How do you generally execute any pre-processing?
      ii. How did you make this decision?
      iii. In which situation(s) might you chose a different method?
      iv. Do you think this step can be automated in the future?
         A. Why (not)?
         B. To what extent?
         C. What are possibilities for automation?

Simulation Period

1. What simulation period do you generally use?
2. How did you make this decision?
3. In which situation(s) might you chose a longer/shorter period?
4. Do you think this step can be automated in the future?
   a) Why (not)?
   b) To what extent?
   c) What are possibilities for automation?

Temporal Resolution

1. What temporal resolution do you generally use?
2. How did you make this decision?
3. In which situation(s) might you chose a different resolution?
4. Do you think this step can be automated in the future?
    a) Why (not)?
    b) To what extent?
    c) What are possibilities for automation?

Spatial Resolution

1. What spatial resolution do you generally use?
2. How did you make this decision?
3. In which situation(s) might you chose a different resolution?
4. Do you think this step can be automated in the future?
    a) Why (not)?
    b) To what extent?
    c) What are possibilities for automation?

Sensitivity Analysis

1. Is this a step you normally execute?
    a) If no:
        i. Why not?
        ii. When would you pre-process the data?
        iii. If this step were to be automated would you use it / execute this step?
            A. Would you find this useful?
            B. Why (not)?

   b) If yes:
      i. How do you generally execute any sensitivity analysis?
      ii. How did you make this decision?
      iii. In which situation(s) might you chose a different method?
      iv. Do you think this step can be automated in the future?
         A. Why (not)?
         B. To what extent?
         C. What are possibilities for automation?

Calibration

1. Do you normally calibrate the model?
   a) If no:
      i. Why not?
      ii. When would you pre-process the data?
      iii. If this step were to be automated would you use it / execute this step?
         A. Would you find this useful?
         B. Why (not)?
   b) If yes:
      i. How do you generally execute any calibration?
      ii. How many parameters do you calibrate on?
      iii. How did you make this decision?
      iv. In which situation(s) might you chose a different method?
      v. Do you think this step can be automated in the future?
         A. Why (not)?
         B. To what extent?
         C. What are possibilities for automation?

Uncertainty Analysis

1. Is this a step you normally execute?
   a) If no:
      i. Why not?

ii. When would you pre-process the data?
iii. If this step were to be automated would you use it / execute this step?
A. Would you find this useful?
B. Why (not)?

b) If yes:

i. How do you generally execute any uncertainty analysis?
ii. How did you make this decision?
iii. In which situation(s) might you chose a different method?
iv. Do you think this step can be automated in the future?
A. Why (not)?
B. To what extent?
C. What are possibilities for automation?

Validation

1. Is this a step you normally execute?

a) If no:

i. Why not?
ii. When would you pre-process the data?
iii. If this step were to be automated would you use it / execute this step?
A. Would you find this useful?
B. Why (not)?

b) If yes:

i. How do you generally validate the model?
ii. How did you make this decision?
iii. In which situation(s) might you chose a different method?
iv. Do you think this step can be automated in the future?
A. Why (not)?
B. To what extent?
C. What are possibilities for automation?

Results/Conclusion

1. How do you reach your final conclusions?
    a) Purely the model?
    b) Multiple model runs?
    c) Expert judgement in combination with the model?

Questions regarding modelling in general:

1. Do you have confidence in a model and its simulations?
2. How doe you estimate your influence as a modeller on the final outcomes?
    a) If another modeller would have executed the same study, would the results be different?
    b) And what about the conclusions?
3. How do you estimate the influence of a modeller on the model in comparison to the programmer?
4. Does your organisation use a certain modelling workflow?
    a) If yes, what does it look like?
    b) What is it based on?
5. Are you familiar with the Dutch handbook 'Good Modelling Practices'? This was published by STOWA.
    a) If no:
        i. What would a handbook be useful for?
        ii. In what format would you use a handbook?
    b) If yes:
        i. Have you used it in the past? Why (not)?
        ii. How could the handbook become more relevant for practical use?
6. Do you have any further remarks or additions?

If it became apparent that automation scripts were already being used, these additional follow-up questions would be asked:

1. What has been automated?
2. What is it based on?
3. How long do you already use it?
4. How is the automation script maintained?

Appendix of Chapter 2: A - ATLAS.ti Codes

The inductive content analysis used in **Chapter 2** resulted in 96 codes. These codes covered the motivations our interviewees discussed. These codes are detailed in Table 1. For each code, the table contains the following characteristics: Code Name, Description, Count and (Motivation) Category. The codes are grouped per motivation category: Individual, Team, Organisational, External, Commissioner, International, National and Consequential. After this, codes regarding Vision are included. At the end, we included multiple codes that were not divided in a category (labelled Miscellaneous).

Table 1 | List of the inductive codes used in the content analysis of **Chapter 2**. The codes are grouped per motivation category, indicated in the last column. This column also specifies in brackets the requirement for the fit-for-purpose framework we aligned this code with. The U stands for Usefulness, R for Reliability and F for Feasibility. If no letter is included, we were unable to group this code in one of the fit-for-purpose requirements. In the third column, the number of occurrences of the codes is detailed.

Code Name	Description	Count	Category
Based on the data	Depending on the data, a modeller makes a choice	9	Individual, Team [R]
Computational time	The total computational time is taken into account when making a decision	59	Individual, Team [F]
Create understanding	To get a (quick) understanding of how the model functions or what the hydrological processes are	7	Individual, Team [R]
Data suitability	If the data is suitable or representative for a certain modelling application	5	Individual [U]
Does not require much attention	According to the modeller, this aspect does not need much attention	1	Individual
Ease of use	How simple it is to implement a method during the modelling process	7	Individual

Code Name	Description	Count	Category
Glad if calibration works	The modeller is glad that the calibration works and is able to obtain results from it	2	Individual
Hardly matters	The modeller thinks there will be no or hardly any difference	13	Individual [F]
Hydrological knowledge / processes	A modeller's knowledge of the hydrological processes	42	Individual, Team [R]
I don't know	The modeller does not know the motivation behind this decision (anymore)	27	Individual
Importance for area	It is deemed necessary for the area the modeller models	5	Individual, Team [R]
Knowledge of area	The knowledge someone has of an area	19	Individual, Team [R]
Logic	By thinking logically something can be determined, i.e. common sense	9	Individual [R]
Most up-to-date	This is the newest version available	13	Individual, Team [F]
No experience	A modeller has no experience, so modellers do not ask for or execute certain parts of the modelling process	8	Individual, Team [F]
No longer a choice	The modeller perceive a certain option as not a choice (anymore)	16	Individual, Standards
Personal expectation	The modeller's expectation of how a decision influences the results	13	Individual
Personal experience	The modeller's experience determines which decision is made	115	Individual [F]
Personal insight	The modeller knows what is needed or is a better option	152	Individual [F]
Personal preference	The modeller's preference for a certain option	63	Individual, Vision

Code Name	Description	Count	Category
Personal trust	The modeller's trust in a certain option influences the decision they make	25	Individual
Set priorities	What is important to do first	2	Individual, Team [F]
Testing	Through testing / trial-and-error make a decision	77	Individual, Team [R]
To try to get better results	To try and get a result that is as accurate as possible	2	Individual, Team [R]
Truthfulness	To try and create a model that is closer to the reality	18	Individual, Team [R]
Best available	At the moment an option was the best available	12	Team, Organisational [F]
Colleagues' preference	The prefered method or option of colleagues	10	Team, Vision
Experience colleague(s)	Based on the experience of colleagues	53	Team [F]
In consultation with superior	A decision is discussed with a superior	1	Team [F]
In consultation with team	A decision is discussed with the team	23	Team [F]
Insight of team	Based on the insight of the team	32	Team [F]
Trust of team	Based on what the team deems trustworthy	8	Team
Vision of team	How the team envisions the modelling process	26	Team, Vision
Board decides	The board of an organisation decides about what needs to be done or used	3	Organisational [U]
Computational power	The computational power available in an organisation	6	Organisational [F]

7

Code Name	Description	Count	Category
Earlier work within organisation	Previously certain work has been executed in the organisation	23	Organisational [F]
Executed by a different department	A different department within an organisation performs certain aspects of the modelling proces	11	Organisational
Internal experience	Experience available within an organisation	7	Organisational [F]
Limit costs	Save costs or take available budget into account	13	Organisational, Commissioner [F]
Necessity	It is deemed necessary to use a certain method (availability, upkeep, etc.)	5	Organisational [F]
Organisational infrastructure	How an organisation has set up their modelling infrastructure impacts a decision	3	Organisational [F]
Partake in development	An organisation is involved in the development of a certain tool or method	4	Organisational, Vision
Sort of obliged	As an organisation they have certain obligations	1	Organisational
Standard in organisation	A general way of working within the organisation	57	Organisational, Standards
Storage capacity	The storage capacity available on a computer or system of an organisation	4	Organisational [F]
Trust of organisation	The trust that is prevalent within an organisation for a certain choice	5	Organisational
Vision organisation	The general vision an organisation has set out	49	Organisational, Vision
Advice from external partner	An external partner advises the modeller	9	External

Code Name	Description	Count	Category
Executed by external partner	An external partner executes (a part) of the modelling process and makes all the decisions relevant to it	37	External
Experience third party	Based on the experience of people outside of your organisation	14	External [F]
External partner uses it	An external partner used this option, which is adopted by the modeller	4	External
External partner determines	An external partner makes the decision for the modeller	6	External
In consultation with third parties	A decision is discussed with parties from other disciplines	3	External [F]
In consultation with external colleagues	A decision is discussed with external colleagues within the hydrological discipline	17	External [F]
Take into account stakeholders/users/policy makers	Taking into account stakeholders, users or policy makers in a decision	4	External [R]
Tender requirements	An external partner details what they want to do in a tender	3	External [U]
Commissioner determines	The commissioner of the modelling study determines what needs to be done or used	45	Commissioner [U]
In consultation with commissioner	The commissioner and contractor discuss what decision is made	15	Commissioner [U]
Provides commissioner	The commissioner provides certain tools, methods or data	12	Commissioner [F]
Requirements for study	Requirements are formulated for a modelling study	13	Commissioner [U]

7

Code Name	Description	Count	Category
Time available in project	The amount of time available within a project	37	Commissioner [F]
What is detailed in tender	Based on the content of quotations, in reply to a public tender	6	Commissioner [U]
Agreed upon nationally	Within the Netherlands certain agreements have been made	2	National [F]
Based on guidelines	A decision is based on the Dutch guideline for Good Modelling Practice	1	National
General standard way of working	A general standard way of working for the whole of the Netherlands	15	National
Collaboration between water authorities	The water authorities have collaborated to determine this/ buy this	10	National [F]
Law	Certain national laws determine which decision is made	4	National [U]
Origin	Where a certain method was developed on a national level	5	National
Widely supported in the Netherlands	All water authorities and consulting companies use and trust it	12	National
International agreements	Internationally certain aspects are prescibed or shared	2	International [F]
Certain format of tools	Earlier chosen tools or methods require a certain format	10	Consequential [F]
Depends on previous choice	Because of an earlier decision, the current decision is influenced	23	Consequential
Determined (automatically) by model	The model (automatically) chooses a certain option	9	Consequential
Evident from previous model step	A certain decision is determined due to the outcomes of a previous modelling step	7	Consequential

Code Name	Description	Count	Category
Happens automatically	A previously written script determines how a modelling step is executed	18	Consequential
Standard value in model	Within a model a standard value is incorporated that is not changed by the modeller	6	Consequential
Consistency	Consistency within an organisation or project	21	Vision [R]
Desired level of detail	The level of detail a modeller envisions for their model or results	36	Vision [U]
Efficiency	The ability to execute something a quickly/easily/accurately as possible	8	Vision [F]
Simplicity	Keep something a simple as possible	10	Vision
Usefulness	The perceived usefulness according to the modeller	14	Vision [U]
Availability	If an option or method is available for the modeller	7	Miscellaneous [F]
Based on theory	Based on hydrological theory	4	Miscellaneous [R]
Characteristics of method	Based on the characteristics of multiple methods a decision is made for one option	28	Miscellaneous [U]
Data availability	If the data is available for the modeller	74	Miscellaneous [F]
Data availability openly	Data are freely available for everyone	6	Miscellaneous [F]
Data quality	What is the data quality?	53	Miscellaneous [F]
Data resolution	The available resolution of the data determines the decision made	9	Miscellaneous [U]
Maintenance is discontinued	The upkeep of a method has been discontinued	4	Miscellaneous [F]

7

Code Name	Description	Count	Category
Model stability	The model needs to be able to run	11	Miscellaneous [F]
New techniques available	New (and better) methods have become available, so those methods are chosen	6	Miscellaneous [F]
Openly available / free	A method is freely available for everyone	7	Miscellaneous [F]
Depends on the area	Dependent on the characteristics of the area certain decisions are made	66	Depends [R]
Depends on the goal	Dependent on the question or goal of the project	132	Depends [U]
Depends on the model	Dependent on the model that is chosen	38	Depends [F]
Depends on the project	Dependent on the project and resources available certain decisions are made	9	Depends [U]
Depends on the situation	Dependent on the circumstances/multiple factors	14	Depends [U]

Appendix of Chapter 3: A - ATLAS.ti Deductive Interview Codes

In Table 2, the interview codes based on Pagano et al. (2016) can be found. These were used for the deductive content analysis in **Chapter 3**.

Table 2 | List of the deductive interview codes used in the content analysis of **Chapter 3**. The deductive interview codes are based on Pagano et al. (2016). In the third column the aspect(s) to which the interview code are linked are given. I1, 2 and 3 stand for respectively the different issues described by Pagano et al. (2016): the role of the modeller, the change in modellers' behaviour and the perception of thrustworthiness. BP1, 2, 3, 4 and 5 correspond to respectively these best practices from Pagano et al. (2016): have transparent systems, no peeking at the answer, evaluate your results, never stop learning the science and redefine your role.

Code Name	Description	Aspects of Pagano et al. (2016)
Need to execute different tasks	Due to the automation, the modeller needs to execute different tasks than they previously did.	I1
New tasks outside own expertise	The modeller's current tasks after automation are outside their own expertise	I1
Different skills obtained by modeller (in future)	Due to automation, the modeller obtains different or more/less skills than they previously did without automation	I1
Different knowledge obtained by modeller (in future)	The modeller obtains different, more or less hydrological knowledge than they previously did without automation	I1
Capacity to take over automation	A modeller would be less capable to take over if the automation fails	I1
Capacity to interpret results	Any change in if the modeller's ability to interpret results due to automation	BP4, BP5
Test of own understanding (give hypothesis)	Before the automation is started, the modeller should give their own hypothesis to test this.	BP2
Bias due to output	The modeller's judgement is influenced by the output they see from the automation, especially if they didn't have an initial idea of the potential outcome	BP2

Code Name	Description	Aspects of Pagano et al. (2016)
Redefinition of the modeller's role	After the automation is incorporated, how is the modeller's role redefined?	BP5
Transfer of who makes the modelling choices	The choices within the modelling process are taken by someone else (often the programmer)	I2
Different set of choices that have to be made by modeller	The modellers is faced with other (modelling) decisions than previously	I1
Change in communication of results	Change in what and how the results are communicated	I2, BP5
(False) 1st impression	The first impression a modeller has of the automation is incorrect	I3
Transference of 1st impression	The first impression of a modeller is copied by other modeller(s)	I3
Change of 1st impression	A modeller changes their view from their initial impression	I3
Transparency of automation process	Clarity of what happens within the automation, i.e. what choices were made? How does the automation process work?	BP1
Obtain intermediate results	While the automation runs, results should be given after different steps to give more insight into the automation.	BP1
Evaluation	Evaluation of the automation, does this happen? How?	BP3, BP4

Appendix of Chapter 3: B - ATLAS.ti Inductive Interview Codes

In Table 3, the interview codes created during the inductive content analysis of **Chapter 3** are shown. The interview codes were subdivided in seven different groups: Extent of automation, Implementation of automation, Interaction between water authorities and consulting company, Levels in automation development, Motivations for (not) developing automation, Role of modeller and programmer and Usage of automation. Some interview codes were not subdivided in the groups. These are in the miscellaneous group.

Table 3 | List of the inductive interview codes used in the content analysis of **Chapter 3**.

Group	Code Name	Description
Extent of automation	Current extent of automation	Indication that the modelling process is automated (to some extent).
	Can be automated	The modelling process can be automated according to the modeller.
	Can't be automated	The modelling process cannot be automated according to the modeller.
	Too extensive	Indication to which extent the modelling process has been automated according to the modeller.
	Completely	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Almost completely	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Mostly	Indication to which extent the modelling process has been or can be automated according to the modeller.
	In development	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Partly	Indication to which extent the modelling process has been or can be automated according to the modeller.

Group	Code Name	Description
Extent of automation (continued)	Not that far	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Not at all	Indication to which extent the modelling process has been or can be automated according to the modeller.
	Executed manually	A certain aspect of the modelling process is carried out manually, without automation.
	Execution of modelling decisions in automation	The automation only covers the steps to execute a certain choice made by the modeller.
	Less complex components	According to the modeller only less complex components are automated.
	No transfer of modelling decision	In automating the modelling process no transfer of modelling decision has occurred for the modeller.
Implementation of automation	Context dependent	How the automation script development is done depends on context (resources, funding, capabilities)
	Dependent on model software	How an automation script is developed depends on with which model software it should be compatible.
	External organisation makes automation script	The automation script development is done by an external organisation.
	First gain insight, then automate	Before an automation script is developed, insight into the process to be automated should exist.
	Iteratively	An automation script is developed over time, improving it whenever it is used.
	Quick and dirty	The automation is developed quickly and for its then current purposes reliable.

Group	Code Name	Description
Implementation of automation (continued)	What is the purpose?	An automation script is developed according to a certain purpose.
Interaction between water authorities and consulting companies	Executed by Consulting company (originally)	Regardless of automation or not, this step is generally executed by a consulting company.
	Executed by Research Insitute	Regardless of automation or not, this step is generally executed by a research insitute.
	Executed by Water Authority (originally)	Regardless of automation or not, this step is generally executed by a water authority.
	Ownership of automation	Who is the owner and therefore responsible for the automation.
Levels in automation development	Organised on personal level	The automation is developed and used at the personal level.
	Organised on team level	The automation is developed and used at the team level.
	Organised on organisational level	The automation is developed and used at the organisational level.
	Organised on inter-organisational level	The automation is developed and used at the inter-organisational level.
	Capacity to automate	What resources (e.g.\ funding and computer capacity) are available to automate within an organisation.
	Create documentation	Along with any automation script, documentation should be written to communicate the purpose of automation script and how it works.
	Differences between organisations	There are differences in modelling practices and perspectives between different organisations.
	Differences within department	There are differences in modelling practices and perspectives within a department.

7

Group	Code Name	Description
Levels in automation development (continued)	Differences within organisation	There are differences in modelling practices and perspectives within an organisation.
	One-time use	
	Open-Source	The automation is available open-source.
	Overlap in tools	Within different departments or organisations similar automation scripts or tools exist.
	Uniformise automation scripts	When scripts from an Individual or Team level are rewritten in such a way that they can be used at a(n) (Inter-) organisational level.
Motivations for (not) developing automation	Automated with previous method	It used to be automated with a previous method, so it is expected that is will be automated in the new method as well.
	Do not execute this ourselves	Modellers do not carry out a certain aspect of the modelling process themselves and therefore, they care less about if and how this aspect is automated.
	Consistency	Automating a certain aspect of the modelling process will create more consistency in the execution of that aspect.
	Easier to do manually	It is considered easier to execute a modelling step manually than to automate the modelling step.
	If needed	Modelling process is automated once it seems necessary.
	Maintain control	The modeller wants to maintain control over the modelling process.
	Not time efficient	Automating a certain part of the modelling process would not be time efficient.

Group	Code Name	Description
Motivations for (not) developing automation (continued)	Not useful	Automating a certain part of the modelling process is deemed not useful.
	Gain insight	Developing and using automation can give the modeller insight into the modelling process, uncertainties and results.
	Ease of use	Automation is easy to use according to the modeller.
	More accurate	Automation will give a more accurate result.
	Objectivity	Automating the modelling process introduces more objectivity into the whole process.
	Reduce human errors	Automation limits the potential for human errors.
	Reproducibility	Automation increases the reproducibility of the modelling results.
	Time efficient	Automation will save time for the modellers.
	Transparency	Automation makes the modelling process more transparent.
	Useful to have	The automation of (a part of) the modelling process is deemed useful.
	Budget	The financial resources to develop an automation.
	Does it matter?	The method or the outcome of the method are not that important.
	Data are limiting factor	Data quality and quantitity limit the potential to develop and use automation.
	Level of difficulty	How difficult the automation is will influence to which extent if at all a certain part of the modelling process is automated.
	Should we want this?	Consideration if the automation of certain parts of the modelling process would be desirably.

7

Group	Code Name	Description
Motivations for (not) developing automation (continued)	Too complex	Automating certain parts of the modelling process would have to many options or be too specific or require the modeller's expertise to execute that part.
Role of modeller and programmer	Programmer most influential	Modeller's perception that programmer is more influential.
	Modelling steps afterwards	The modelling steps taken after the programmer has developed an automation script are more influential than the decisions made in the automation script.
	Programmer follows hydrological laws	The modeller's trust that the programmer follows hydrological laws.
	Modeller's responsibility	What is considered to be the modeller's responsibility in using and developing an automation script.
	Programmer's responsibility	What is considered to be the programmer's responsibility in using and developing an automation script.
Usage of automation	Was not used	A previous automation script was not used.
	Would not use it	A modeller would not use a certain automation script when automated.
	Would execute this modelling step when automated	If a certain modelling step were to be automated, a modeller would use it in the future.
	Not sure if they'd use it	If a certain aspect is automated, a modeller is not sure they would use it.
	Couple of times	An automation script would be used a couple of times immediately after development.
	Irregular	The use of the automation script would be / is irregular.
	Repeated use	The automation script would be / is used frequently.

Group	Code Name	Description
Usage of automation (continued)	Use default value	Within an automation script the default value was used even if there was a choice.
	Check by modeller	The modeller would check the output after the automation has given it.
	Modeller gives some input	When using an automation script, a modeller will give some inputs to the automation script before it runs.
	Make modelling decision yourself	The modeller want to make the modelling decision themselves.
	Expert knowledge necessary	The experience and expertise of a modeller is necessary in a particular modelling step.
	Remain critical	When using an automation script, a modeller should keep checking the automation itself.
	Understand the automation	When using an automation script, the modeller should understand the automation.
	Use automation to advise you	Use the results of an automated aspect to inform you as a modeller to make a next decision and maybe adapt certain aspects of the step just executed with the automation script.
Miscellaneous	Data doesn't receive enough attention	A modeller perceives that the data are payed enough attention.
	No choice	A modeller does not have a choice in a particular modelling step, which makes automation illogical or difficult.
	Not automation, but rather standardisation	A modeller perceives that standardisation is essential, not automation necessarily.
	Not known by interviewee	The interviewee does not know (exactly).

7

Group	Code Name	Description
Miscellaneous (continued)	Surprised not organised on inter-organisational level	A modeller is surprised that a certain part of the modelling process is not developed or used at an inter-organisational level.
	Trust in scientific literature high	A modellers does not trust certain aspects of science with regards to the modelling process or automation.
	Trust in scientific literature low	A modellers trusts certain aspects of the modelling process or automation because they are underpinned by science.

Appendix of Chapter 4: A - Interview guides

Interview guide - Australia

The interview guide consists of three parts: one about the background, one about the specifics of the water management and modelling in the Murray-Darling Basin and the eWater Source framework and at the end some final questions. The interview guide included here is the generic one we adapted depending on the interview. For some interviews the adaptations were quite substantial, for others less so. The adaptations were made depending on the time available during the interview and the expertise of the interviewee.

Questions regarding the background:

1. How would you describe your current position?
2. What is your background?
    a) In hydrological modelling?
    b) In work experience?
3. How would you describe your work in the Murray-Darling basin?
    a) Current work?
    b) Past work?
4. Can you describe your experience as a modeller?
    a) Which type of models?
    b) Which models exactly?
    c) How many years?

Specific questions:

Questions about their current modelling practice:

1. What does your current modelling process look like?
    a) Which modelling software?
    b) What are the decisions you make?
    c) Does a guideline exist (on institutional, team or personal level)?
2. What did your modelling process look like prior to 2007?
    a) Which modelling software?
    b) What are the decisions you make?
    c) Does a guideline exist (on institutional, team or personal level)?

Questions about Basin Plan:

1. Are you familiar with the Basin Plan?
    a) If yes, to what extent?
    b) If no:
        i. Why not?
        ii. Would you familiarise yourself in the future?
            A. Why (not)?
2. What changes to the modelling process have occurred since the basin plan was formed?
    a) Did anything change to the modelling process?
    b) Which modelling steps were adapted?
    c) How were these adapted?
3. How do you perceive the basin plan?
4. Are you familiar with your state's Water Resource Plan (WRP)?
    a) If yes, to what extent?

  b) If no:
    i. Why not?
    ii. Would you familiarise yourself in the future?
      A. Why (not)?

5. Did the WRPs have an effect on the modelling process?
  a) If yes:
    i. How?
    ii. What changed?
  b) If no:
    i. Why not?

6. How do you perceive the WRPs?

7. Are you familiar with the Uplift Programme?
  a) If yes, to what extent?
  b) If no:
    i. Why not?
    ii. Would you familiarise yourself in the future?
      A. Why (not)?

8. Did the Uplift Programme have an impact on your modelling process?
  a) If yes:
    i. How?
    ii. What changed?
  b) If no:
    i. Why not?
    ii. Would you familiarise yourself in the future?
      A. Why (not)?

9. How do you perceive the Uplift Programme?

10. What do you deem necessary in these efforts to harmonise the modelling in the MDB?
    a) What would you need?
    b) What would you prefer?

Final questions:

1. Do you have any recommendations or referrals for other people that would be good candidates for my study?
2. Do you have any remarks or additions?

Interview guide - The Netherlands developers

The interview guide has three parts: one about the background of the interviewee, one about the process of creating and publishing the guideline and one regarding some final general interview questions.

Questions regarding the background:

1. How would you describe your current position?
2. How would you describe your position when the Dutch guideline was written?
3. Can you describe your experience as a modeller?
    a) What type of modelling studies?
    b) How many years? Then and now?

Questions regarding the guideline:

Initiation/Idea of guideline

1. How did the initiative for creating the guideline originate?
    a) By whom?
    b) What was the initiative?

   c) When?
   d) Why?
   e) How?

2. How did the organisation you worked at get involved?

3. How did you get involved with the development of the guideline?

4. The Standaard Raamwerk Water has been created around the same time, how do these two initiatives relate to each other?

   a) And how did it influence each other?
   b) Is it still used? Has it ever been used?

5. On February 4 1997, there was a meeting. Multiple parties were involved.

   a) Who were involved?
   b) How did these parties get involved?
   c) How was communication handled with parties that were uninvolved?

Process of creating the guideline

1. After the initiative was taken, how was the guideline created?

   a) By whom?
   b) What was the initiative?
   c) When?
   d) How?

2. In the guideline, a test phase after the first draft is described. How did this go?

   a) Who were involved?
   b) How was the test group selected?
   c) What were the outcomes of this test phase?
   d) How were these outcomes processed?

After publication:

1. How was the guideline disclosed?
    a) Certain communication?
    b) Any promotion?
2. How was the guideline picked up after publication?
    a) Was there a difference between disciplines? Or commissioners?
    b) Who did the guideline reach? Modellers? Commissioners?
    c) Was there a difference between who was or was not involved in the guideline's development?
    d)
3. Within the guideline, certain goals are set. Have these goals been reached?
    a) "To provide a starting point for principles regarding model use that are supported by all involved parties in the water management"
    b) "To initiate more careful use of models in water management"
    c) "To improve the reproducibility and transferability of model studies"
4. Within the guideline, it is described that the guideline should be kept 'dynamic'. How did this turn out?
    a) Test phase? How was this set up?
        i. Who were involved?
        ii. What were the outcomes of this?
    b) Were there any submission for update suggestions or requests?
        i. How was this spread out over time?
        ii. Would you tackle this differently now?
    c) How were any updates to the guideline communicated?

Reflection:

1. How do you reflect on this whole process?
   a) Are there aspects you would do differently? If yes, what?
   b) Do you have any suggestions if someone were to create a guideline today?
2. Are there any adaptations you would make to the guideline?

Other standardisation initiative:

1. Around the same time as the development of the guideline, other initiatives were also developed, for example Standaard Raamwerk Water, Adventusstelsel and CIW standaard. How do these different initiatives relate?
2. Are there other initiatives that I am not aware of? Or certain projects that were executed?

Final questions:

1. Do you have any recommendations or referrals for other people that would be good candidates for my study?
2. Do you have any remarks or additions?

Appendix of Chapter 4: B - ATLAS.ti Codes

Table 4 contains the codes from the inductive content analysis of **Chapter 4**. The codes were categorised in seven axial codes: Development, Application, Advantages, Disadvantages, Improvements, Risks and Updates.

Table 4 | List of the inductive codes from our inductive content analysis.

Axial Code	Inductive Code	Description	Count
Development	Agreements made	Agreements were made on the standardisation approach.	3
	Automation	Previous automation influences the standardisation approach and automation occurs after a standardisation approach is implemented.	3
	Based on experience	The standardisation approach is based on the experience of modellers.	10
	Comparability	Through developing a standardisation approach, the aim is to increase comparability of modelling studies.	7
	Compatibility	Through developing a standardisation approach, the aim is to increase compatibility between modelling studies.	4
	Cooperatively	The standardisation approach is developed and implemented in collaboration. (Also grouped in axial code Application)	49
	Create consistency	Through developing a standardisation approach, the aim is to increase consistency in the modelling process.	42

Axial Code	Inductive Code	Description	Count
Development (continued)	Create more efficiency	Through developing a standardisation approach, the aim is to increase efficiency of the modelling process and within organisations.	4
	Create support base for policy decisions	Through developing a standardisation approach, the aim is to create support for policy decisions.	33
	Create transparency	Through developing a standardisation approach, the aim is to increase transparency of the modelling process.	24
	Crises driven development	The standardisation approach's development is initiated or accelerated due to a natural crisis.	12
	Different modelling practices	Between states, organisations and people, different modelling practices exist. (Also grouped in axial code Development)	60
	Driven by other considerations	The development of a standardisation approach is motivated by other external factors.	2
	Each state had their own model	Each state had their own legacy models, which they had used for years.	30
	General consensus	A consensus needs to be reached to choose a standardisation approach.	12
	Inconsistency in basin	Within the basin, a lot of inconsistency occurs, both in nature and in the modelling process. (Also grouped in axial code Disadvantages)	13
	Increase collaboration	Through developing a standardisation approach, the aim is to increase between stakeholders.	7
	Inter-modeller variability	Through developing a standardisation approach, the aim is to decrease inter-modeller variability.	17

7

Axial Code	Inductive Code	Description	Count
Development (continued)	Iteratively	The standardisation approach is developed iteratively.	3
	Meet requirements	The modelling process and workflow is designed based on the requirements that need to be met.	24
	Models need to be run for legislation	Models are run to inform and possible provide numbers for legislation.	11
	Needs to withstand (public) scrutiny	Model results need to withstand (public) scrutiny.	16
	Negotiation	Negotation is necessary to reach a consensus on what the standardisation approach will be.	39
	No reproducibility	Through developing a standardisation approach, the aim is to increase reproducibility of modelling studies.	1
	Originated organically	The standardisation approach developed organically.	9
	Political pressures	Political influences have an impact on the standardisation approach and the modelling. (Also grouped in axial code Disadvantages)	34
	Robustness	Through developing a standardisation approach, the aim is to increase robustness of modelling studies.	2
	Selection of who is involved	A selection is made on who is involved in the development of the standardisation approach.	12
	States' independence	The states have their own independence concerning their water management and modelling. (Also grouped in axial code Development)	10

Axial Code	Inductive Code	Description	Count
Development (continued)	Top down	The standardisation approach is developed and implemented top-down.	25
	Vision		10
Application	Apply policy changes in model	Policy changes are incorporated in the model software.	25
	As reference	Guidelines can be used as a reference for (junior) modellers. (Also grouped in axial code Advantages)	21
	Bigger than anticipated	The development and implementation of the standardisation approach was more work than anticipated.	2
	Build guidelines	To support models or new model software and to provide more consistency in modelling practices, guidelines can be created	8
	Changed relationship between stakeholders	The standardisation approach changes the relationship between stakeholders.	23
	Communication	Communication between stakeholders about how the standardisation approach is developed and implemented and promoted.	28
	Community driven implementation	The implementation of the standardisation approach depends on the community.	9
	Cooperatively	The standardisation approach is developed and implemented in collaboration. (Also grouped in axial code Development)	49
	Creates tension	The development or the implementation of the standardisation approach creates tension between stakeholders.	14

7

Axial Code	Inductive Code	Description	Count
Application (continued)	Familiarity	Modellers tend to use what they are familiar with. (Also grouped in axial code Disadvantages)	10
	How (often) used?	Indication of how and how often the interviewees used the modelling approach.	41
	How standardisation can be implemented	Suggestions and experiences on how to the standardisation approach was implemented.	45
	Initially dealing with bugs	In the beginning, users of the standardisation approach had to deal with bugs.	8
	Initially done, but less and less over time	The standardisation approach was initially used but the usage decreased over time.	2
	Legacy	Modelling practices and models have a legacy in its development and usage in certain organisations, states or for certain people.	18
	Lot of effort put in	A lot of effort has already been put into developing and implementing the standardisation approach.	1
	Ongoing	The development and implementation of the standardisation approach is still ongoing.	16
	Policy influenced by modelling	Model results or the modelling process influences the policy that is made.	13
	Policy situation is complicating	The policy situation complicates development, implementation and usage of the standardisation approach. (Also grouped in axial code Application)	28

Axial Code	Inductive Code	Description	Count
Application (continued)	Provide incentive	In order to implement the standardisation approach an incentive can be provided to move it along. (Also grouped in axial code Application)	5
	Provide training	A training is provided to teach modellers how to use a new tool or method.	6
	Slow implementation of new methodology	The new methodology or standardisation approach is implemented slowly and carefully	47
	Suitability	Considerations of how a model can be suitable for a certain modelling study or not.	16
	Use best available	Modelling practice is to use the best available.	9
	Use community	The implementation of the standardisation approach relies on the community	11
	Voluntary	The standardisation approach is implemented on voluntary basis.	5
Advantages	Adaptability	The standardisation approach has the potential to evolve over time and to be changed. (Also grouped in axial code Improvements)	5
	As reference	Guidelines can be used as a reference for (junior) modellers. (Also grouped in axial code Application)	20
	Commitment to transition	The stakeholders have a commitment to transition to the new modelling approach.	14
	Create support base for model results	Standardisation can provide a support base for model results.	9
	Idea behind it is good	The idea behind the standardisation approach is good.	5

7

Axial Code	Inductive Code	Description	Count
Advantages (continued)	Improve connection between stakeholders	The standardisation approach has brought the modelling community closer together, but it can still improve engagement and relations between stakeholders further. (Also grouped in axial code Improvements)	20
	Improve flexibility	The tools already have some flexibility, but this can be enhanced. (Also grouped in axial code Improvements)	10
	Improves quality of modelling	Standardisation improves the quality of the modelling process.	18
	Increases communication between models	Standardisation eases the ability to link different models.	12
	Legitimacy	Standardisation can provide legitimacy to model results in the decision-support domain.	2
	More accessibility	Standardisation can provide more accessibility to models, tools or data.	8
	More collaboration	Standardisation can result in more collaboration between stakeholders.	6
	More consistency	Standardisation can result in more consistency between organisations, modellers and modelling studies.	23
	More credibility	Standardisation can provide more credibility to model results in the decision-support domain.	2
	More efficiency	Standardisation can result in more efficiency of the modelling process or within organisations.	18
	More resources	Standardisation can result in more resources being available.	12

Axial Code	Inductive Code	Description	Count
Advantages (continued)	More robust	Standardisation can result in more robustness of the modelling studies.	5
	More transparency	Standardisation can result in more transparency of the model results and modelling process.	12
	More understanding	Standardisation can create more understanding of the hydrological processes and the modelling process and the model results.	8
	Practices changed due to collaboration	Modelling practices changed due to collaboration (within the standardisation approach).	16
	Reproducibility	Standardisation can result in more reproducibility of the modelling studies.	14
	Retraceability	Standardisation can result in more retraceability of the modelling study.	3
	Seems useful	The standardisation approach seems to be useful.	7
	Shaped thinking of (future) modellers	The development of the standardisation approach shaped the thinking of (future) modellers.	2
	Showing technical capabilities	The standardisation approach showed the capabilities of another method.	7
	Successful	The standardisation approach has been (quite) successful (given its limitations/resources).	12
	Transferability	The models and also the knowledge obtained by modellers is more easily transferred.	13
Disadvantages	Accessibility limited	The access to the tool, model or data is limited.	8

Axial Code	Inductive Code	Description	Count
Disadvantages (continued)	Capability limited	The knowledge and know-how is not present to be able to implement or use the standardisation approach or model.	19
	Capacity limited	The available resources or humanpower restricts the usage or development of the standardisation approach.	39
	Communication is difficult	Communication hinders the implementation or usage of the standardisation approach.	9
	Complex	The model or standardisation approach is complex, impeding its usage.	17
	Costly	The costs of the standardisation approach are high.	20
	Different modelling practices	Between states, organisations and people, different modelling practices exist. (Also grouped in axial code Development)	45
	Different priorities	The priorities between institutes differ, resulting in different opinion on what should be prioritised.	40
	Different terminology	A different terminology of the modelling process or hydrological processes exists between institutes and states.	9
	Does not meet requirements	The standardisation approach or the model does not meet the requirements.	3
	Every modelling study is specific	Each modelling study has to be tailored to the subject that is studied.	26
	Everything for everybody	The standardisation approach has been developed in such a way that it accommodates everyone's old modelling practices.	17

Axial Code	Inductive Code	Description	Count
Disadvantages (continued)	Familiarity	Modellers tend to use what they are familiar with. (Also grouped in axial code Application)	10
	First user's practice captured	In the standardisation approach, they captured the first user experience.	7
	Generic	The standardisation approach is too generic.	10
	Improve documentation	Documentation of the models and modelling workflow can still be improved. (Also grouped in axial code Improvements)	8
	Inconsistency in basin	Within the basin, a lot of inconsistency occurs, both in nature and in the modelling process. (Also grouped in axial code Development)	13
	Less efficiency	The standardisation approach has introduced less efficiency.	1
	Less of a community than wanted	With the implementation of the standardisation approach, less of a community was built than initially envisioned.	3
	Less transparency	The standardisation approach has resulted in less transparency than before.	10
	Lock-in effect	Due to the standardisation approach, a lock-in effect is happening.	2
	Loss of power	Due to the standardisation approach a loss of power can be experienced by any stakeholder. (Also grouped in axial code Risks)	3
	Model isn't capable to model reality	The model is not able to represent the processes occurring in the world.	3

7

Axial Code	Inductive Code	Description	Count
Disadvantages (continued)	Modelling practice unchanged	Even though the standardisation approach is implemented, the actual modelling practices remain unchanged.	15
	Not successful	The standardisation approach is considered to be unsuccessful (in certain aspects).	11
	Not user friendly	The standardisation approach is not user friendly.	19
	Not very useful	The standardisation approach is not deemed very useful.	4
	Numbers are different	The standardisation approach results in different results from the models.	25
	One general standardisation approach not possible	It is considered that one general standardisation approach is not possible to create.	9
	Policy situation is complicating	The policy situation complicates development, implementation and usage of the standardisation approach. (Also grouped in axial code Application)	28
	Political pressures	Political influences have an impact on the standardisation approach and the modelling. (Also grouped in axial code Development)	34
	Slower software	The new software is slower than the previous software.	13
	States' independence	Implementing a standardisation approach need to take into account that the states have their own independence concerning their water management and modelling. (Also grouped in axial code Development)	10

Axial Code	Inductive Code	Description	Count
Improvements	Adaptability	The standardisation approach has the potential to evolve over time and to be changed. (Also grouped in axial code Advantages)	5
	Ambitious timeline	The improvements scheduled on an ambitious timeline.	3
	Build capability	Increasing knowledge and know-how of the models and tools used, also with regards to evolving and using the standardisation approach.	12
	Build capacity	Increasing the resources and humanpower to be able to implement and update the standardisation approach.	8
	Build trust	Increase stakeholders' trust in models and their results, the standardisation approach and each other.	34
	Could be improved	The standardisation approach can be improved further.	3
	Evolving standardisation approach	Suggestions for how the standardisation approach can become more usable and relevant, both in its format, but also technically.	68
	Funding for maintenance	Concerns and possibilities regarding the resources available for the upkeep of the standardisation approach. (Also grouped in axial code Risks)	22
	Improve connection between stakeholders	The standardisation approach has brought the modelling community closer together, but it can still improve engagement and relations between stakeholders further. (Also grouped in axial code Advantages)	20

7

Axial Code	Inductive Code	Description	Count
Improvements (continued)	Improve documentation	Documentation of the models and modelling workflow can still be improved. (Also grouped in axial code Disadvantages)	8
	Improve ease of use	The usability of the tools and standardisation approaches can still be increased.	7
	Improve flexibility	The tools already have some flexibility, but this can be enhanced. (Also grouped in axial code Advantages)	10
	Improve modelling practice	The modelling practice currently used can still be improved, for example to create more consistency.	25
	Include processes	Within the modelling tools, certain processes are not (yet) represented (sufficiently).	22
	Increase knowledge base	The science behind the modelling should be developed further again.	32
	Need collaboration	For the standardisation approach to be implemented collaboration is needed.	1
	Need consistency	Improvements are needed to the standardisation approach to create more consistency.	3
	Not necessary to improve	It is deemed unnecessary (yet) to improve the standardisation approach or tool.	2
	Ownership is difficult	Arranging ownership of a standardisation approach or model can be challenging.	7
	Provide incentive	In order to implement the standardisation approach an incentive can be provided to move it along. (Also grouped in axial code Application)	5

Axial Code	Inductive Code	Description	Count
Risks	Funding for maintenance	Concerns and possibilities regarding the resources available for the upkeep of the standardisation approach. (Also grouped in axial code Improvements)	22
	Longevity	Concerns with regards to how long a standardisation approach can continue.	4
	Loss of power	Due to the standardisation approach a loss of power can be experienced by any stakeholder. (Also grouped in axial code Disadvantages)	3
	No longer usable	The developed tool is no longer usable.	1
Updates	Audit models	New models used for decision support are reviewed before they are used.	16
	Easier maintenance	Updates to the standardisation approach mean that its upkeep will be easier.	2
	Further developed internationally	The standardisation approach was picked up internationally, where it was developed further.	3
	No evaluation	There is no evaluation of the standardisation approach or modelling process.	9
	No requirements to review models	The review of models is on a voluntary basis, there is no obligation to do so.	4
	Not enough time	There is not enough time available to properly review or evaluate the modelling processes.	2
	Review (modelling) process	The executed (modelling) process is evaluated.	13
	Review upcoming	At the time of the interviews, a review was scheduled to be executed.	4

7

Axial Code	Inductive Code	Description	Count
Updates (continued)	Small pool of reviewers	There are only few people available to execute the necessary reviews.	3
	Under review	At the time of the interviews, a review was being executed.	5
	Update models	Models are kept up-to-date.	8
	Updated ad-hoc	Models and the standardisation approach are updated when necessary.	8

Bibliography

Addor, N. & L. Melsen (2019). Legacy, rather than adequacy, drives the selection of hydrological models. *Water Resources Research*, 55(1), 378–390. doi: 10.1029/2018WR022958.

Adecs Airinfra Consultants (2021). Berekeningen lelystad airport - achtergrondrapport bij de passende beoordeling.

Aitsi-Selmi, A., K. Blanchard & V. Murray (2016). Ensuring science is useful, usable and used in global disaster risk reduction and sustainable development: a view through the sendai framework lens. *Palgrave Communications*, 2(1), 1–9. doi: 10.1057/palcomms.2016.16.

Alder, K. & N. Wise (1995). A revolution to measure. *The values of precision*, 4. doi: 10.1515/9780691218120-004.

Amorim, L. D. & J. Cai (2015). Modelling recurrent events: a tutorial for analysis in epidemiology. *International journal of epidemiology*, 44(1), 324–333. doi: 10.1093/ije/dyu222.

Anderson, D. (2014). *Endurance: Australian stories of drought*. Csiro publishing.

Andrienko, N., G. Andrienko, S. Chen & B. Fisher (2022). Seeking patterns of visual pattern discovery for knowledge building. In *Computer Graphics Forum*, volume 41, pages 124–148. Wiley Online Library. doi: 10.1111/cgf.14515.

ATLAS.ti Scientific Software Development GmbH (2022). ATLAS.ti Version 9.1.7.0, Retrieved from https://atlasti.com.

Babel, L., D. Vinck & D. Karssenberg (2019). Decision-making in model construction: unveiling habits. *Environmental Modelling & Software*. doi: 10.1016/j.envsoft.2019.07.015.

Bacchelli, A. & C. Bird (2013). Expectations, outcomes, and challenges of modern code review. In *2013 35th International Conference on Software Engineering (ICSE)*, pages 712–721. IEEE. doi: 10.1109/ICSE.2013.6606617.

Balci, O. (2001). A methodology for certification of modeling and simulation applications. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 11(4), 352–377. doi: 10.1145/508366.508369.

Baldissera Pacchetti, M., J. Jebeile & E. Thompson (2024). For a pluralism of climate modeling strategies. *Bulletin of the American Meteorological Society*, 105(7), E1350–E1364. doi: 10.1175/BAMS-D-23-0169.1.

Barnett, B., L. Townley, V. Post, R. Evans, R. Hunt, L. Peeters, S. Richardson, A. Werner, A. Knapton & A. Boronkay (2012). *Australian Groundwater Modelling Guidelines*.

Bartholomeus, R. P., K. van der Wiel, A. F. van Loon, M. H. van Huijgevoort, M. T. van Vliet, M. Mens, S. Muurling-van Geffen, N. Wanders & W. Pot (2023). Managing water across the flood–drought spectrum: Experiences from and challenges for the netherlands. *Cambridge Prisms: Water*, 1, e2. doi: 10.1017/wat.2023.4.

Bartram, D. & R. Bayliss (1984). Automated testing: Past, present and future. *Journal of Occupational Psychology*, 57(3), 221–237. doi: 10.1111/j.2044-8325.1984.tb00164.x.

Bastin, L., D. Cornford, R. Jones, G. B. Heuvelink, E. Pebesma, C. Stasch, S. Nativi, P. Mazzetti & M. Williams (2013). Managing uncertainty in integrated environmental modelling: The uncertweb framework. *Environmental Modelling & Software*, 39, 116–134. doi: 10.1016/j.envsoft.2012.02.008.

Beck, C. & G. T. Jones (1989). Bias and archaeological classification. *American Antiquity*, 54(2), 244–262. doi: 10.2307/281706.

Beck, M. & T. Krueger (2016). The epistemic, ethical, and political dimensions of uncertainty in integrated assessment modeling. *Wiley Interdisciplinary Reviews: Climate Change*, 7(5), 627–645. DOI: 10.1002/wcc.415.

Beeston, L., R. Blewitt, S. Bulmer & J. Wilson (2021). Traffic Modelling Guidelines v4. *Transport for London: London, UK.*

Bereiter, C. & M. Scardamalia (2014). *Knowledge Building and Knowledge Creation: One Concept, Two Hills to Climb*, pages 35–52. Springer Singapore, Singapore. DOI: 10.1007/978-981-287-047-6_3.

Berg, M. (1997). Problems and promises of the protocol. *Social science & medicine*, 44(8), 1081–1088. DOI: 10.1016/S0277-9536(96)00235-3.

Beven, K. (2007). Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. *Hydrology and Earth System Sciences*, 11(1), 460–467. DOI: 10.5194/hess-11-460-2007.

Beven, K. & J. Freer (2001). Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the glue methodology. *Journal of hydrology*, 249(1-4), 11–29.

Beven, K. J. (2000). Uniqueness of place and process representations in hydrological modelling. *Hydrology and earth system sciences*, 4(2), 203–213. DOI: 10.5194/hess-4-203-2000.

Bilcke, J., P. Beutels, M. Brisson & M. Jit (2011). Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. *Medical Decision Making*, 31(4), 675–692. DOI: 10.1177/0272989X11409240.

Black, M. (2019). *Models and metaphors: Studies in language and philosophy*. Cornell University Press.

Blackett, P., E. Le Heron, S. Awatere, R. Le Heron, J. Logie, J. Hyslop, J. Ellis, F. Stephenson & J. Hewitt (2024). Navigating choppy waters: why are we always arguing about risk and uncertainty in marine multi-use environments and what can we do about it? *Policy Quarterly*, 20(3), 62–68. DOI: 10.26686/pq.v20i3.9560.

Boaz, A. & C. Hayden (2002). Pro-active evaluators: enabling research to be useful, usable and used. *Evaluation*, 8(4), 440–453. DOI: 10.1177/13563890260620630.

Borowski, I. & M. Hare (2007). Exploring the gap between water managers and researchers: difficulties of model-based tools to support practical water management. *Water Resources Management*, 21, 1049–1074. DOI: 10.1007/s11269-006-9098-z.

Brouwer, S. (2015). *Dutch Water Management*, pages 73–85. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-17241-5_3.

Burt, T. P. & J. J. McDonnell (2015). Whither field hydrology? the need for discovery science and outrageous hydrological hypotheses. *Water Resources Research*, 51(8), 5919–5928. DOI: 10.1002/2014WR016839.

Butts, M. B., J. T. Payne, M. Kristensen & H. Madsen (2004). An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation. *Journal of Hydrology*, 298(1-4), 242–266. DOI: 10.1016/j.jhydrol.2004.03.042.

Calder, M., C. Craig, D. Culley, R. de Cani, C. A. Donnelly, R. Douglas, B. Edmonds, J. Gascoigne, N. Gilbert, C. Hargrove, D. Hinds, D. C. Lane, D. Mitchell, G. Pavey, D. Robertson, B. Rosewell, S. Sherwin, M. Walport & A. Wilson (2018). Computational modelling for decision-making: where, why, what, who and how. *Royal Society open science*, 5(6), 172096. DOI: 10.1098/rsos.172096.

Cambridge Dictionary (2024a). Meaning of Standard in English. Https://dictionary.cambridge.org/dictionary/english/standard, Last accessed February 13 2024.

7

Cambridge Dictionary (2024b). Meaning of Vision in English. Https://dictionary.cambridge.org/dictionary/english/vision, Last accessed February 13 2024.

Cambridge Dictionary (SA). Meaning of Standardisation in English. Https://dictionary.cambridge.org/dictionary/english/standardization, Last accessed May 3 2024.

Carlisle, M., C. Gimbar & J.G. Jenkins (2023). Auditor-client interactions—an exploration of power dynamics during audit evidence collection. *Auditing: A Journal of Practice & Theory*, 42(1), 27–51. doi: 10.2308/AJPT-2020-130.

Caro, J. J., A. H. Briggs, U. Siebert & K. M. Kuntz (2012). Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. *Medical Decision Making*, 32(5), 667–677. doi: 10.1177/0272989X12454577.

Carver, L. & M. Turoff (2007). Human-computer interaction: the human and computer as a team in emergency management information systems. *Communications of the ACM*, 50(3), 33–38. doi: 10.1145/1226736.1226761.

Cash, D., W.C. Clark, F. Alcock, N.M. Dickson, N. Eckley & J. Jäger (2002). Salience, credibility, legitimacy and boundaries: linking research, assessment and decision making. *Assessment and Decision Making (November 2002)*. doi: 10.2139/ssrn.372280.

Clark, M.P., A.G. Slater, D.E. Rupp, R.A. Woods, J.A. Vrugt, H.V. Gupta, T. Wagener & L.E. Hay (2008). Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, 44(12). doi: 10.1029/2007WR006735.

Cockerill, K., P. Glynn, I. Chabay, M. Farooque, R. Hämäläinen, B. Miyamoto & P. McKay (2019). Records of engagement and decision making for environmental and socio-ecological challenges. *EURO Journal on decision processes*, 7(3-4), 243–265. doi: 10.1007/s40070-019-00104-6.

Cole, S. (1983). The hierarchy of the sciences? *American Journal of sociology*, 89(1), 111–139. doi: 10.1086/227835.

Comte, A. (1855). *The positive philosophy of Auguste Comte*. C. Blanchard.

Connell, D. & R.Q. Grafton (2011). Water reform in the murray-darling basin. *Water Resources Research*, 47(12). doi: 10.1029/2010WR009820.

Cosgrove, W.J. & D.P. Loucks (2015). Water management: Current and future challenges and research directions. *Water Resources Research*, 51(6), 4823–4839. doi: 10.1002/2014WR016869.

Craig, J.R., G. Brown, R. Chlumsky, W. Jenkinson, G. Jost, K. Lee, J. Mai, M. Serrer, N. Sgro, M. Shafii, A.P. Snowdon & B.A. Tolson (2020). Flexible watershed simulation with the raven hydrological modelling framework. *Environmental Modelling & Software*, page 104728. doi: 10.1016/j.envsoft.2020.104728.

Craik, K.J.W. (1967). *The nature of explanation*, volume 445. CUP Archive.

Crout, N., T. Kokkonen, A. Jakeman, J. Norton, L. Newham, R. Anderson, H. Assaf, B. Croke, N. Gaber, J. Gibbons, D. Holzworth, J. Mysiak, J. Reichl, R. Seppelt, T. Wagener & P. Whitfield (2008). Chapter two good modelling practice. *Developments in Integrated Environmental Assessment*, 3, 15–31. doi: 10.1016/S1574-101X(08)00602-9.

Daly, C., W.P. Gibson, G.H. Taylor, M.K. Doggett & J.I. Smith (2007). Observer bias in daily precipitation measurements at United States cooperative network stations. *Bulletin of the American Meteorological Society*, 88(6), 899–912. doi: 10.1175/BAMS-88-6-899.

Daniels, G. & T. Starr (1997). Guidelines for funding operations and maintenance of intelligent transportation systems/advanced traffic management systems. *Transportation research record*, 1588(1), 49–62. doi:

10.3141/1588-07.

Darling, J. (1945). Speaking of flood control. *Journal (American Water Works Association)*, 37(4), 345–349.

Daston, L. & P. L. Galison (2007). *Objectivity*. Princeton University Press.

de Campos Vallim, C. H. & A. T. Akabane (2023). Detection and control of traffic jams in urban environment. In *Brazilian Technology Symposium*, pages 60–74. Springer. doi: 10.1007/978-3-031-66961-3_6.

Deitrick, A. R., S. A. Torhan & C. A. Grady (2021). Investigating the influence of ethical and epistemic values on decisions in the watershed modeling process. *Water Resources Research*, 57(12), e2021WR030481. doi: 10.1016/j.ijdrr.2018.10.023.

Deltares (2023a). D-HYDRO Suite 1D2D. Https://www.deltares.nl/software-en-data/producten/d-hydro-suite-1d2d, Last accessed July 6 2023.

Deltares (2023b). HYDROLIB. Https://deltares.github.io/HYDROLIB/, Last accessed July 6 2023.

Deltares (2023c). HYDROLIB github. Https://github.com/Deltares/HYDROLIB, Last accessed July 6 2023.

Deltares (2023d). Sobek Suite. Https://www.deltares.nl/en/software/sobek/, Last accessed July 6 2023.

Demortain, D. (2008). Standardising through concepts: The power of scientific experts in international standard-setting. *Science and Public Policy*, 35(6), 391–402. doi: 10.3152/030234208X323325.

Devi, G. K., B. P. Ganasri & G. S. Dwarakish (2015). A review on hydrological models. *Aquatic procedia*, 4, 1001–1007. doi: 10.1016/j.aqpro.2015.02.126.

Doorn, N. (2017). Resilience indicators: Opportunities for including distributive justice concerns in disaster management. *Journal of Risk Research*, 20(6), 711–731. doi: 10.1080/13669877.2015.1100662.

Doorn, N. (2021). Artificial intelligence in the water domain: Opportunities for responsible use. *Science of the Total Environment*, 755, 142561. doi: 10.1016/j.scitotenv.2020.142561.

Dupre, L. (1967). Situation ethics and objective morality. *Theological studies*, 28(2), 245–257. doi: 10.1177/004056396702800202.

Dutch Water Authorities (SA). About. Https://dutchwaterauthorities.com/about/, Last accessed on June 18 2024.

Eilander, D., H. Boisgontier, L. J. e. Bouaziz, J. Buitink, A. Couasnon, B. Dalmijn, M. Hegnauer, T. de Jong, S. Loos, I. Marth & W. van Verseveld (2023). HydroMT: Automated and reproducible model building and analysis. *Journal of Open Source Software*, 8(83), 4897. doi: 10.21105/joss.04897.

Emerson, R. W. (2015). Convenience sampling, random sampling, and snowball sampling: How does sampling affect the validity of research? *Journal of visual impairment & blindness*, 109(2), 164–168. doi: 10.1177/0145482X1510900215.

Evanoff, R. J. (2004). Universalist, relativist, and constructivist approaches to intercultural ethics. *International Journal of Intercultural Relations*, 28(5), 439–458. doi: https://doi.org/10.1016/j.ijintrel.2004.08.002.

eWater Source Group (SA). eWater Source. Https://ewater.org.au/ewater-solutions/tools/source/, Last accessed on August 10 2024.

Eyring, V., P. M. Cox, G. M. Flato, P. J. Gleckler, G. Abramowitz, P. Caldwell, W. D. Collins, B. K. Gier, A. D. Hall, F. M. Hoffman, G. C. Hurtt, A. Jahn, C. D. Jones, S. A. Klein, J. P. Krasting, L. Kwiatkowski, R. Lorenz, E. Maloney, G. A. Meehl, A. G. Pendergrass, R. Pincus, A. C. Ruane, J. L. Russel, B. M. Sanderson, B. D. Santer, S. C. Sherwood, R. J. Stouffer & M. S. Williamson (2019). Taking climate model evaluation to the next level. *Nature Climate Change*, page 1. doi:

10.1038/s41558-018-0355-y.

Fan, S. L., C. Y. Lee, H. Y. Chong & M. J. Skibniewski (2018). A critical review of legal issues and solutions associated with building information modelling. *Technological and Economic Development of Economy*, 24(5), 2098–2130. doi: 10.3846/tede.2018.5695.

Fanelli, D. (2010). "positive" results increase down the hierarchy of the sciences. *PloS one*, 5(4), e10068. doi: 10.1371/journal.pone.0010068.

Fenicia, F., D. Kavetski & H. H. G. Savenije (2011). Elements of a flexible approach for conceptual hydrological modeling: 1. motivation and theoretical development. *Water Resources Research*, 47(W11510). doi: 10.1029/2010WR010174.

Fleming, S. (2009). An informal survey of watershed model users: Preferences, applications, and rationales. *Streamline Watershed Management Bulletin*, 13(1), 32–35.

Frigg, R. & S. Hartmann (2024). Models in Science. In E. N. Zalta & U. Nodelman (editors), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Fall 2024 edition.

Funtowicz, S. O. & J. R. Ravetz (1993). Science for the post-normal age. *Futures*, 25(7), 739–755. doi: 10.1016/00163287(93)90022-L.

Garner, M. & S. Hamilton (2011). Principles of epidemiological modelling. *Revue Scientifique et Technique-OIE*, 30(2), 407.

Glattfelder, J. B. (2019). *Philosophy and Science: What Can I Know?*, pages 299–344. Springer International Publishing, Cham. doi: 10.1007/978-3-030-03633-1_9.

Glynn, P. D., A. A. Voinov, C. D. Shapiro & P. A. White (2017). From data to decisions: Processing information, biases, and beliefs for improved management of natural resources and environments. *Earth's Future*, 5(4), 356–378. doi: 10.1002/2016EF000487.

Gnaden, D. & S. Holdaway (2000). Understanding observer variation when recording stone artifacts. *American antiquity*, 65(4), 739–747. doi: 10.2307/2694425.

Goodenough, A. E., W. S. Carpenter, L. McTavish, B. Blades, E. Clarke, S. Griffiths, N. Harding, R. Scott, E. Walsh, L. Wilson & A. G. Hart (2020). The impact of inter-observer variability on the accuracy, precision and utility of a commonly-used grassland condition index. *Ecological indicators*, 117, 106664. doi: 10.1016/j.ecolind.2020.106664.

Goulden, S., M. Negev, S. Reicher & T. Berman (2019). Implications of standards in setting environmental policy. *Environmental science & policy*, 98, 39–46. doi: 10.1016/j.envsci.2019.05.002.

Government of the Netherlands (SA). Water management in The Netherlands. Https://www.government.nl/topics/water-management/water-management-in-the-netherlands, Last accessed on July 27 2023.

Grafton, R. Q. & J. Horne (2014). *Global Water: Issues and Insights*, chapter Water markets in the Murray-Darling basin, pages 37 – 44. ANU Press.

Grimm, V. (2020). The odd protocol: an update with guidance to support wider and more consistent use. *Ecological Modelling*, 428, 109105. doi: j.ecolmodel.2020.109105.

Grimm, V., J. Augusiak, A. Focks, B. M. Frank, F. Gabsi, A. S. Johnston, C. Liu, B. T. Martin, M. Meli, V. Radchuk, P. Thorbek & S. F. Railsback (2014). Towards better modelling and decision support: Documenting model development, testing, and analysis using TRACE. *Ecological modelling*, 280, 129–139. doi: 10.1016/j.ecolmodel.2014.01.018.

Guba, E. G. & Y. S. Lincoln (1994). Competing paradigms in qualitative research. *Handbook of qualitative research*, 2(163-194), 105.

Gundersen, O. E. (2021). The fundamental principles of reproducibility. *Philosophical Transactions of the Royal Society A*, 379(2197), 20200210. doi: 10.1098/rsta.2020.0210.

Gustafsson, N. (1981). *A Review of Methods for Objective Analysis*, pages 17–76. Springer New York, New York, NY. doi: 10.1007/978-1-4612-5970-1_2.

Hamilton, S. H., B. Fu, J. H. Guillaume, J. Badham, S. Elsawah, P. Gober, R. J. Hunt, T. Iwanaga, A. J. Jakeman, D. P. Ames, A. Curtis, M. C. Hill, S. A. Pierce & F. Zare (2019). A framework for characterising and evaluating the effectiveness of environmental modelling. *Environmental modelling & software*, 118, 83–98. doi: 10.1016/j.envsoft.2019.04.008.

Hamilton, S. H., C. A. Pollino, D. S. Stratford, B. Fu & A. J. Jakeman (2022). Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility. *Environmental Modelling & Software*, 148, 105278. doi: 10.1016/j.envsoft.2021.105278.

Haraway, D. (2013). Situated knowledges: The science question in feminism and the privilege of partial perspective 1. In *Women, science, and technology*, pages 455–472. Routledge.

Hart, B. T. (2016). The australian murray-darling basin plan: factors leading to its successful development. *Ecohydrology & Hydrobiology*, 16(4), 229–241. doi: 10.1016/j.ecohyd.2016.09.002.

Hegselmann, R., U. Mueller & K. G. Troitzsch (1996). *Modelling and simulation in the social sciences from the philosophy of science point of view*, volume 23. Springer Science & Business Media.

Henckens, G. & W. Engel (2017). *Benchmark inundatiemodellen: modelfunctionaliteiten en testbank berekeningen*. STOWA. Https://3diwatermanagement.com/uploads/sites/5/STOWA-2017-34.pdf, Last accessed on October 14 2024.

Henriksen, H. J., L. Troldborg, P. Nyegaard, T. O. Sonnenborg, J. C. Refsgaard & B. Madsen (2003). Methodology for construction, calibration and validation of a national hydrological model for Denmark. *Journal of Hydrology*, 280(1-4), 52–71. doi: 10.1016/S0022-1694(03)00186-0.

Holländer, H. M., T. Blume, H. Bormann, W. Buytaert, G. B. Chirico, J. F. Exbrayat, D. Gustafsson, H. Hölzel, P. Kraft, C. H. Stamm, G. Blöschl & H. Flühler (2009). Comparative predictions of discharge from an artificial catchment (Chicken Creek) using sparse data. *Hydrology and Earth System Sciences*, 13(11), 2069–2094. doi: 10.5194/hess-13-2069-2009.

Holländer, H. M., H. Bormann, T. Blume, W. Buytaert, G. B. Chirico, J. F. Exbrayat, D. Gustafsson, H. Hölzel, T. Krausse, P. Kraft, C. H. Stamm, G. Blöschl & H. Flühler (2014). Impact of modellers' decisions on hydrological a priori predictions. *Hydrology and Earth System Sciences*, 18(6), 2065–2085. doi: 10.5194/hess-18-2065-2014.

Hoogewoud, J., G. Prinsen, J. Hunink, A. Veldhuizen, F. van der Bolt & W. de Lange (2013). *Toetsingsrapportage NHI 3.0*. Deltares, Netherlands. Project 1206107-000.

Horton, P., B. Schaefli & M. Kauzlaric (2022). Why do we have so many different hydrological models? a review based on the case of switzerland. *Wiley Interdisciplinary Reviews: Water*, 9(1), e1574.

Howard, R. & B. C. Björk (2008). Building information modelling–Experts' views on standardisation and industry deployment. *Advanced engineering informatics*, 22(2), 271–280. doi: 10.1016/j.aei.2007.03.001.

Hughes, D., D. Kingston & M. Todd (2011). Uncertainty in water resources availability in the okavango river basin as a result of climate change. *Hydrology and Earth System Sciences*, 15(3), 931–941. doi: 10.5194/hess-15-931-2011.

Irani, M., M. M. Naderi, A. R. M. Bavani, E. Hassanzadeh & H. Moftakhari (2024). A framework for coastal flood hazard assessment under sea level rise: Application to the persian gulf. *Journal of Environ-*

mental Management, 349, 119502. DOI: 10.1016/j.jenvman.2023.119502.

Jakeman, A. J., S. Elsawah, H. H. Wang, S. H. Hamilton, L. Melsen & V. Grimm (2024). Towards normalizing good practice across the whole modeling cycle: its instrumentation and future research topics. *Socio-Environmental Systems Modelling*, 6, 18755–18755. DOI: 10.18174/sesmo.18755.

Jakeman, A. J., R. A. Letcher & J. P. Norton (2006). Ten iterative steps in development and evaluation of environmental models. *Environmental Modelling & Software*, 21(5), 602–614. DOI: 10.1016/j.envsoft.2006.01.004.

Jimenez, A. V. (2022). *Standards and standardisation: the social construction of uniformity*. Ph.D. thesis, Macquarie University.

Jones, H. F., D. Özkundakci, S. Hunt, H. Giles & B. Jenkins (2020). Bridging the gap: A strategic framework for implementing best practice guidelines in environmental modelling. *Environmental Science & Policy*, 114, 533–541. DOI: 10.1016/j.envsci.2020.09.030.

Joskowicz, L., D. Cohen, N. Caplan & J. Sosna (2019). Inter-observer variability of manual contour delineation of structures in CT. *European radiology*, 29, 1391–1399. DOI: 10.1007/s00330-018-5695-5.

Kajko-Mattsson, M. (2005). A survey of documentation practice within corrective maintenance. *Empirical Software Engineering*, 10, 31–55. DOI: 10.1023/B:LIDA.0000048322.42751.ca.

Kassahun, A., H. Scholten, G. Zompanakis & C. Gavadinas (2005). Support for model based water management with the harmoniqua toolbox. In *International Conference on Complexity and Integrated Resources Management, Osnabrück, Germany, 14-17 June 2004*, pages 1282–1287. Citeseer.

Kim, J., Y. Cui, E. Hwang, D. Franklin & Y. Seo (2020). "012012 or 111000": preference for consumption pattern-seeking. *European Journal of Marketing*, 54(9), 2171–2194. DOI: 10.1108/EJM-03-2019-0224.

Kirby, M., R. Bark, J. Connor, M. E. Qureshi & S. Keyworth (2014). Sustainable irrigation: How did irrigated agriculture in australia's murray–darling basin adapt in the millennium drought? *Agricultural Water Management*, 145, 154–162. DOI: 10.1016/j.agwat.2014.02.013.

Klein, A., K. Unverzagt, R. Alba, J. F. Donges, T. Hertz, T. Krueger, E. Lindkvist, R. Martin, J. Niewöhner, H. Prawitz, M. Schlüter, L. Schwarz & N. Wijermans (2024). From situated knowledges to situated modelling: a relational framework for simulation modelling. *Ecosystems and People*, 20(1), 2361706. DOI: 10.1080/26395916.2024.2361706.

Knutti, R., D. Masson & A. Gettelman (2013). Climate model genealogy: Generation CMIP5 and how we got there. *Geophysical Research Letters*, 40(6), 1194–1199. DOI: 10.1002/grl.50256.

Kok, M., K. Slager, H. de Moel, W. Botzen, K. de Bruijn, D. Wagenaar, S. Rikkert, E. Koks & K. van Ginkel (2023). Rapid damage assessment caused by the flooding event 2021 in limburg, netherlands. *Journal of Coastal and Riverine Flood Risk*, 2. DOI: 10.59490/jcrfr.2023.0010.

Kopp, R. E., E. A. Gilmore, C. M. Little, J. Lorenzo-Trueba, V. C. Ramenzoni & W. V. Sweet (2019). Usable science for managing the risks of sea-level rise. *Earth's future*, 7(12), 1235–1269. DOI: 10.1029/2018EF001145.

Krueger, T. & R. Alba (2022). Ontological and epistemological commitments in interdisciplinary water research: Uncertainty as an entry point for reflexion. *Frontiers in Water*, 4, 1038322. DOI: 10.3389/frwa.2022.1038322.

Krueger, T., C. Maynard, G. Carr, A. Bruns, E. N. Mueller & S. Lane (2016). A transdisciplinary account of water research. *Wiley Interdisciplinary Reviews: Water*, 3(3), 369–389. DOI: 10.1002/wat2.1132.

Krueger, T., T. Page, K. Hubacek, L. Smith & K. Hiscock (2012). The role of expert opinion in environmental modelling. *Environmental Modelling & Software*, 36, 4–18. DOI: 10.1016/j.envsoft.2012.01.011.

Kukuric, N. (2011). Nationaal hydrologisch instrumentarium.

Lahtinen, T. J., J. H. Guillaume & R. P. Hämäläinen (2017). Why pay attention to paths in the practice of environmental modelling? *Environmental Modelling & Software*, 92, 74–81. doi: 10.1016/j.envsoft.2017.02.019.

Lane, S. N. (2014). Acting, predicting and intervening in a socio-hydrological world. *Hydrology and Earth System Sciences*, 18(3), 927–952. doi: 10.5194/hess-18-927-2014.

de Lange, W., J. Hoogewoud, G. Prinsen, J. Verkaik & A. Veldhuizen (2015). De stand van zaken rond het nationaal hydrologisch instrumentarium. *H2O: tijdschrift voor watervoorziening en afvalwaterbehandeling*, 45(9), 17–19.

Laplane, L., P. Mantovani, R. Adolphs, H. Chang, A. Mantovani, M. McFall-Ngai, C. Rovelli, E. Sober & T. Pradeu (2019). Why science needs philosophy. *Proceedings of the National Academy of Sciences*, 116(10), 3948–3952. doi: 10.1073/pnas.1900357116.

Latour, B. (1990). Technology is society made durable. *The sociological review*, 38(S1), 103–131. doi: 10.1111/j.1467-954X.1990.tb03350.x.

Law, J. (2004). *After method: Mess in social science research*. Routledge. doi: 10.4324/9780203481141.

Lazo, J. K., R. E. Morss & J. L. Demuth (2009). 300 billion served: Sources, perceptions, uses, and values of weather forecasts. *Bulletin of the American Meteorological Society*, 90(6), 785–798. doi: 10.1175/2008BAMS2604.1.

Leamer, E. E. (2009). *Introduction: We Are Pattern-Seeking, Story-Telling Animals*, pages 3–16. Springer Berlin Heidelberg, Berlin, Heidelberg. doi: 10.1007/978-3-540-46389-4_1.

Lenhard, J. & E. Winsberg (2010). Holism, entrenchment, and the future of climate model pluralism. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 41(3), 253–262. doi: 10.1016/j.shpsb.2010.07.001.

Levine, J. R. & W. J. Wilson (2013). Poverty, politics, and a "circle of promise": Holistic education policy in boston and the challenge of institutional entrenchment. *Journal of urban affairs*, 35(1), 7–24. doi: 10.1111/juaf.12001.

Lewis, E., S. Birkinshaw, C. Kilsby & H. J. Fowler (2018). Development of a system for automated setup of a physically-based, spatially-distributed hydrological model for catchments in Great Britain. *Environmental modelling & software*, 108, 102–110. doi: 10.1016/j.envsoft.2018.07.006.

Lim, T. C., P. D. Glynn, G. W. Shenk, P. Bitterman, J. H. Guillaume, J. C. Little & D. Webster (2023). Recognizing political influences in participatory social-ecological systems modeling. *Socio-Environmental Systems Modelling*, 5, 18509–18509. doi: 10.18174/sesmo.18509.

Limerick, H., D. Coyle & J. W. Moore (2014). The experience of agency in human-computer interactions: a review. *Frontiers in human neuroscience*, 8, 643. doi: 10.3389/fnhum.2014.00643.

Lin, A. M. (2015). Researcher positionality. *Research methods in language policy and planning: A practical guide*, pages 21–32. doi: 10.1002/9781118340349.ch3.

Longino, H. E. (1990). *Science as social knowledge: Values and objectivity in scientific inquiry*. Princeton university press.

Lord, E. & S. Anthony (2000). MAGPIE: A modelling framework for evaluating nitrate losses at national and catchment scales. *Soil Use and Management*, 16, 167–174. doi: 10.1111/j.1475-2743.2000.tb00222.x.

Maier, H. R., F. R. Taghikhah, E. Nabavi, S. Razavi, H. Gupta, W. Wu, D. A. Radford & J. Huang (2024). How much x is in xai: Responsible use of "explainable" artificial intelligence in hydrology and water resources. *Journal of Hydrology X*, page 100185. DOI: 10.1016/j.hydroa.2024.100185.

Malano, H. (2010). Modelling and decision making in water resource management. *IAHS Publication*, 338, 111–126.

Mayer, L. A., K. Loa, B. Cwik, N. Tuana, K. Keller, C. Gonnerman, A. M. Parker & R. J. Lempert (2017). Understanding scientists' computational modeling decisions about climate risk management strategies using values-informed mental models. *Global Environmental Change*, 42, 107–116. DOI: 10.1016/j.gloenvcha.2016.12.007.

McCracken, S., S. E. Salterio & M. Gibbins (2008). Auditor–client management relationships and roles in negotiating financial reporting. *Accounting, organizations and society*, 33(4-5), 362–383. DOI: 10.1016/j.aos.2007.09.002.

McFarlane, D. C. & K. A. Latorella (2002). The scope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction*, 17(1), 1–61. DOI: 10.1207/S15327051HCI1701_1.

McMillan, H., J. Seibert, A. Petersen-Overleir, M. Lang, P. White, T. Snelder, K. Rutherford, T. Krueger, R. Mason & J. Kiang (2017). How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications. *Water Resources Research*, 53(7), 5220–5228. DOI: 10.1002/2016WR020328.

McMillan, H. K., I. K. Westerberg & T. Krueger (2018). Hydrological data uncertainty and its implications. *Wiley Interdisciplinary Reviews: Water*, 5(6), e1319. DOI: 10.1002/wat2.1319.

MDBA (SA). Types of water modelling. https://www.mdba.gov.au/water-management/water-resource-modelling/types-water-modelling, Last accessed on August 22 2024.

MDBA, M. D. B. A. (2012). Hydrologic modelling to inform the proposed basin plan: Methods and results. https://www.mdba.gov.au/sites/default/files/publications/Hydrologic-Modelling-Report.pdf, Last accessed on September 25 2024.

Meijer, A., M. Wiarda, N. Doorn & G. van de Kaa (2023). Towards responsible standardisation: investigating the importance of responsible innovation for standards development. *Technology Analysis & Strategic Management*, pages 1–15. DOI: 10.1080/09537325.2023.2225108.

Melsen, L. (2022). It takes a village to run a model—the social practices of hydrological modeling. *Water Resources Research*, 58(2), e2021WR030600. DOI: 10.1029/2021WR030600.

Melsen, L. A., N. Addor, N. Mizukami, A. J. Newman, P. J. Torfs, M. P. Clark, R. Uijlenhoet & A. J. Teuling (2018a). Mapping (dis)agreement in hydrologic projections. *Hydrology and Earth System Sciences*, 22(3), 1775–1791. DOI: 10.5194/hess-22-1775-2018.

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

Melsen, L. A., J. Vos & R. Boelens (2018b). What is the role of the model in socio-hydrology? discussion of "Prediction in a socio-hydrological world". *Hydrological Sciences Journal*, 63(9), 1435–1443. DOI: 10.1080/02626667.2018.1499025.

Mendelsohn, E. (1977). The social construction of scientific knowledge. In *The social production of scientific knowledge*, pages 3–26. Springer. DOI: 10.1007/978-94-010-1186-0_1.

Mer, F., W. Baethgen & R. W. Vervoort (2020). Building trust in SWAT model scenarios through a multi-institutional approach in Uruguay. *Socio-Environmental Systems Modelling*, 2, 17892–17892. DOI: 10.18174/

sesmo.2020a17892.

Mercieca, D. & D.P. Mercieca (2013). Engagement with research: acknowledging uncertainty in methodology. *International Journal of Research & Method in Education*, 36(3), 228–240. doi: 10.1080/1743727X.2013.806470.

Merleau-Ponty, M. (1945). Phénoménologie de la perception, gallimard. *Paris: Bibliothèque des idées*.

Micheletti, T., M.C. Wimmler, U. Berger, V. Grimm & E.J. McIntire (2024). Beyond guides, protocols and acronyms: Adoption of good modelling practices depends on challenging academia's status quo in ecology. *Ecological Modelling*, 496, 110829. doi: 10.1016/j.ecolmodel.2024.110829.

Mir, M. & M. Casadesús (2011). Standardised innovation management systems: A case study of the Spanish Standard UNE 166002: 2006. *Innovar*, 21(40), 171–188.

Moallemi, E.A., F. Zare, P.M. Reed, S. Elsawah, M.J. Ryan & B.A. Bryan (2020). Structuring and evaluating decision support processes to enhance the robustness of complex human–natural systems. *Environmental Modelling & Software*, 123, 104551. doi: 10.1016/j.envsoft.2019.104551.

Molle, F. (2009). Water, politics and river basin governance: Repoliticizing approaches to river basin management. *Water International*, 34(1), 62–70. doi: 10.1080/02508060802677846.

Moon, K. & D. Blackman (2014). A guide to understanding social science research for natural scientists. *Conservation biology*, 28(5), 1167–1177. doi: 10.1111/cobi.12326.

Moray, N. (1999). *Mental models in theory and practice*. MIT Press. doi: 10.7551/mitpress/1480.003.0014.

Morrison, L. (2016). Observer error in vegetation surveys: a review. *Journal of Plant Ecology*, 9, 367–379. doi: 10.1093/jpe/rtv077.

Müller, B., S. Balbi, C.M. Buchmann, L. De Sousa, G. Dressler, J. Groeneveld, C.J. Klassert, Q.B. Le, J.D. Millington, H. Nolzen, D.C. Parker, J.G. Polhill, M. Schlüter, J. Schulze, N. Schwarz, Z. Sun, P. Taillandier & H. Weise (2014). Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling & Software*, 55, 156–163. doi: 10.1016/j.envsoft.2014.01.029.

Munteanu, C., H. Molyneaux, W. Moncur, M. Romero, S. O'Donnell & J. Vines (2015). Situational ethics: Re-thinking approaches to formal ethics requirements for human-computer interaction. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 105–114. doi: 10.1145/2702123.2702481.

Nabavi, E. (2018). Failed policies, falling aquifers: Unpacking groundwater overabstraction in iran. *Water Alternatives*, 11(3), 699.

Nabavi, E. (2022). Computing and modeling after covid-19: More responsible, less technical. *IEEE Transactions on Technology and Society*, 3(4), 252–261. doi: 10.1109/TTS.2022.3218738.

Nabavi, E., R. Nicholls & G. Roussos (2024). Locating responsibility in the future of human–ai interactions. *IEEE Transactions on Technology and Society*, 5(1), 58–60. doi: 10.1109/TTS.2024.3386247.

Nearing, G.S., Y. Tian, H.V. Gupta, M.P. Clark, K.W. Harrison & S.V. Weijs (2016). A philosophical basis for hydrological uncertainty. *Hydrological Sciences Journal*, 61(9), 1666–1678.

Nederlands Hydrologisch Instrumentarium (SA). HyDAMO Data Model. Https://nhi.nu/ontwikkelingen/hydamo/, Last accessed on July 30 2023.

Nederlandse Rijksoverheid (SA). Normen voor producten, diensten of bedrijfsprocessen. Https://www.rijksoverheid.nl/onderwerpen/certificaten-keurmerken-en-meetinstrumenten/normen-voor-producten-diensten-of-processen, Last accessed on June 18 2024.

7

Nielsen, J. (1995). *Getting Usability Used*, pages 3–12. Springer US, Boston, MA. doi: 10.1007/978-1-5041-2896-4_1.

Njeri, S. (2021). Race, positionality and the researcher. *The companion to peace and conflict fieldwork*, pages 381–394. doi: 10.1007/978-3-030-46433-2_26.

Nonaka, I. & R. Toyama (2003). The knowledge-creating theory revisited: knowledge creation as a synthesizing process. *Knowledge management research & practice*, 1(1), 2–10. doi: 10.1057/palgrave.kmrp.8500001.

NOS Nieuws (2022). Om onderzoekt mogelijk gesjoemel bij berekening stikstofuitstoot lelystad airport.

Oldfield, M. (2022). Towards pedagogy supporting ethics in modelling. *Journal of Humanistic Mathematics*, 12(2), 128–159. doi: 10.5642/jhummath.XVSP3245.

OMF (SA). The open modeling foundation.

Oreskes, N. (1998). Evaluation (not validation) of quantitative models. *Environmental health perspectives*, 106(suppl 6), 1453–1460. doi: 10.1289/ehp.98106s61453.

Oreskes, N. (2003). The role of quantitative models in science. In *Models in ecosystem science*, volume 13, page 27. Princeton University Press Princeton, NJ.

Oreskes, N. (2018). Why believe a computer? models, measures, and meaning in the natural world. In *The earth around us*, pages 70–82. Routledge.

Otterpohl, R., M. Grottker & J. Lange (1997). Sustainable water and waste management in urban areas. *Water Science and Technology*, 35(9), 121–133. doi: 10.1016/S0273-1223(97)00190-X.

Packett, E., N. J. Grigg, J. Wu, S. M. Cuddy, P. J. Wallbrink & A. J. Jakeman (2020). Mainstreaming gender into water management modelling processes. *Environmental Modelling & Software*, 127, 104683. doi: 10.1016/j.envsoft.2020.104683.

Padilla, J. J., S. Y. Diallo, C. J. Lynch & R. Gore (2018). Observations on the practice and profession of modeling and simulation: A survey approach. *Simulation*, 94(6), 493–506. doi: 10.1177/003754971773715.

Pagano, T. C., F. Pappenberger, A. W. Wood, M. H. Ramos, A. Persson & B. Anderson (2016). Automation and human expertise in operational river forecasting. *Wiley Interdisciplinary Reviews: Water*, 3(5), 692–705. doi: 10.1002/wat2.1163.

Parasuraman, R., T. B. Sheridan & C. D. Wickens (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on systems, man, and cybernetics-Part A: Systems and Humans*, 30(3), 286–297. doi: 10.1109/3468.844354.

Parker, W. S. & E. Winsberg (2018). Values and evidence: how models make a difference. *European Journal for Philosophy of Science*, 8, 125–142. doi: 10.1007/s13194-017-0180-6.

Parliament of Australia (SA). Murray-darling basin water issues. https://www.aph.gov.au/About_Parliament/Parliamentary_Departments/Parliamentary_Library/pubs/BriefingBook43p/murraydarlingissues, Last accessed August 22 2024.

Parnas, D. L. (2010). Precise documentation: The key to better software. In *The Future of Software Engineering*, pages 125–148. Springer. doi: 10.1007/978-3-642-15187-3_8.

Pascarella, L., D. Spadini, F. Palomba, M. Bruntink & A. Bacchelli (2018). Information needs in contemporary code review. *Proceedings of the ACM on human-computer interaction*, 2(CSCW), 1–27. doi: 10.1145/3274404.

Perrin, C., C. Michel & V. Andréassian (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of hydrology*, 279(1-4), 275–289. DOI: 10.1016/S0022-1694(03)00225-7.

Petersen, A. C., A. Cath, M. Hage, E. Kunseler & J. P. van der Sluijs (2011). Post-normal science in practice at the netherlands environmental assessment agency. *Science, Technology, & Human Values*, 36(3), 362–388. DOI: 10.1177/0162243910385797.

Philips, Z., L. Bojke, M. Sculpher, K. Claxton & S. Golder (2006). Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. *Pharmacoeconomics*, 24, 355–371. DOI: 10.2165/00019053-200624040-00006.

Polhill, J. G. & B. Edmonds (2007). Open access for social simulation. *Journal of Artificial Societies and Social Simulation*, 10(3), 10–10.

Proske, U., S. Ferrachat, S. Klampt, M. Abeling & U. Lohmann (2023). Addressing complexity in global aerosol climate model cloud microphysics. *Journal of Advances in Modeling Earth Systems*, 15(5), e2022MS003571. DOI: 10.1029/2022MS003571.

Proske, U., S. Ferrachat, D. Neubauer, M. Staab & U. Lohmann (2022). Assessing the potential for simplification in global climate model cloud microphysics. *Atmospheric Chemistry and Physics*, 22(7), 4737–4762. DOI: 10.5194/acp-22-4737-2022.

Pulkkinen, K., S. Undorf, F. Bender, P. Wikman-Svahn, F. Doblas-Reyes, C. Flynn, G.C. Hegerl, A. Jönsson, G.K. Leung, J. Roussos, T.G. Shepherd & E. Thompson (2022). The value of values in climate science. *Nature Climate Change*, 12(1), 4–6. DOI: 10.1038/s41558-021-01238-9.

Puy, A., M. Massimi, B. Lankford & A. Saltelli (2023). Irrigation modelling needs better epistemology. *Nature Reviews Earth & Environment*, 4(7), 427–428. DOI: 10.1038/s43017-023-00459-0.

Quinn, J. D., P.M. Reed, M. Giuliani & A. Castelletti (2017). Rival framings: A framework for discovering how problem formulation uncertainties shape risk management trade-offs in water resources systems. *Water Resources Research*, 53(8), 7208–7233. DOI: 10.1002/2017WR020524.

Raskin, J. D. & A. E. Debany (2018). The inescapability of ethics and the impossibility of "anything goes": A constructivist model of ethical meaning making. *Journal of Constructivist Psychology*, 31(4), 343–360. DOI: 10.1080/10720537.2017.1383954.

Reed, M. S., A. Graves, N. Dandy, H. Posthumus, K. Hubacek, J. Morris, C. Prell, C. H. Quinn & L. C. Stringer (2009). Who's in and why? a typology of stakeholder analysis methods for natural resource management. *Journal of environmental management*, 90(5), 1933–1949. DOI: 10.1016/j.jenvman.2009.01.001.

Refsgaard, J. C. (1996). *Terminology, Modelling Protocol And Classification of Hydrological Model Codes*, pages 17–39. Springer Netherlands, Dordrecht. DOI: 10.1007/978-94-009-0257-2_2.

Refsgaard, J. C. & H. J. Henriksen (2004). Modelling guidelines—-terminology and guiding principles. *Advances in Water Resources*, 27(1), 71–82. DOI: 10.1016/j.advwatres.2003.08.006.

Refsgaard, J. C., J. P. van der Sluijs, A. L. Højberg & P. A. Vanrolleghem (2007). Uncertainty in the environmental modelling process–a framework and guidance. *Environmental Modelling & Software*, 22(11), 1543–1556. DOI: 10.1016/j.envsoft.2007.02.004.

Remmers, J. O. (2024). Data for project 'Modellers or Influencers? The social aspects of hydrological modelling'. DOI: 10.17026/SS/AXTDZT.

Remmers, J. O., A. J. Teuling & L. A. Melsen (2020). Can model structure families be inferred from model output? *Environmental Modelling & Software*, 133, 104817. DOI: 10.1016/j.envsoft.2020.104817.

Rödder, S., M. Heymann & B. Stevens (2020). Historical, philosophical, and sociological perspectives on earth system modeling. *Journal of Advances in Modeling Earth Systems*, 12(10), e2020MS002139. DOI:

10.1029/2020MS002139.

Ross, A. & H. Chang (2020). Socio-hydrology with hydrosocial theory: two sides of the same coin? *Hydrological Sciences Journal*, 65(9), 1443–1457. doi: 10.1080/02626667.2020.1761023.

Saltelli, A., L. Benini, S. Funtowicz, M. Giampietro, M. Kaiser, E. Reinert & J. P. van der Sluijs (2020). The technique is never neutral. How methodological choices condition the generation of narratives for sustainability. *Environmental Science & Policy*, 106, 87–98. doi: 10.1016/j.envsci.2020.01.008.

Saltelli, A. & M. Di Fiore (2023). *The politics of modelling: Numbers between science and policy*. Oxford University Press. doi: 10.1093/oso/9780198872412.001.0001.

Samaniego, L., A. Bárdossy & R. Kumar (2010a). Streamflow prediction in ungauged catchments using copula-based dissimilarity measures. *Water Resources Research*, 46(2). doi: 10.1029/2008WR007695.

Samaniego, L., R. Kumar & S. Attinger (2010b). Multiscale parameter regionalization of a grid-based hydrologic model at the mesoscale. *Water Resources Research*, 46(5). doi: 10.1029/2008WR007327.

Sampaio, P., P. Saraiva & A. Guimarães Rodrigues (2011). ISO 9001 certification forecasting models. *International Journal of Quality & Reliability Management*, 28(1), 5–26. doi: 10.1108/02656711111097526.

Sanz, D., J. Vos, F. Rambags, J. Hoogesteger, E. Cassiraga & J. J. Gómez-Alday (2019). The social construction and consequences of groundwater modelling: insight from the mancha oriental aquifer, spain. *International Journal of Water Resources Development*, 35(5), 808–829. doi: 10.1080/07900627.2018.1495619.

Savenije, H. H. (2009). Hess opinions" the art of hydrology". *Hydrology and Earth System Sciences*, 13(2), 157–161. doi: 10.5194/hess-13-157-2009.

Schmolke, A., P. Thorbek, D. L. DeAngelis & V. Grimm (2010). Ecological models supporting environmental decision making: a strategy for the future. *Trends in ecology & evolution*, 25(8), 479–486. doi: 10.1016/j.tree.2010.05.001.

Scholten, H., J. Refsgaard & A. Kassahun (2004). Structuring multidisciplinary knowledge for model based water management: the harmoniqua approach. In *Proceedings of iEMSs*, pages 1288–1293.

Schwarz, N., G. Dressler, K. Frank, W. Jager, M. A. Janssen, B. Müller, M. Schlüter, N. Wijermans & J. Groeneveld (2020). Formalising theories of human decision-making for agent-based modelling of social-ecological systems: practical lessons learned and ways forward. *Available at SSRN 4152673*. doi: 10.18174/sesmo.2020a16340.

Seth, J. (1991). Standardisation and quality assurance. In *Principles and practice of immunoassay*, pages 154–189. Springer.

Sharp, G., K. D. Fritscher, V. Pekar, M. Peroni, N. Shusharina, H. Veeraraghavan & J. Yang (2014). Vision 20/20: perspectives on automated image segmentation for radiotherapy. *Medical physics*, 41(5), 050902. doi: 10.1118/1.4871620.

Silver, N. (2012). *The signal and the noise: the art and science of prediction*. Penguin UK.

Simonton, D. K. (2006). Scientific status of disciplines, individuals, and ideas: Empirical analyses of the potential impact of theory. *Review of General Psychology*, 10(2), 98–112. doi: 10.1037/1089-2680.10.2.98.

Sismondo, S. (2011). *An introduction to science and technology studies*. John Wiley & Sons.

Sivapalan, M., H. H. Savenije & G. Blöschl (2012). Socio-hydrology: A new science of people and water. *Hydrol. Process*, 26(8), 1270–1276. doi: 10.1002/hyp.8426.

Smith, N. (1908). Subjectivism and realism in modern philosophy. *The Philosophical Review*, 17(2), 138–148. doi: 10.2307/2177394.

Soedirgo, J. & A. Glas (2020). Toward active reflexivity: Positionality and practice in the production of knowledge. *PS: Political Science & Politics*, 53(3), 527–531. DOI: 10.1017/S1049096519002233.

Sørensen, J., M. Klee, T. Palshof & H. Hansen (1993). Performance status assessment in cancer patients. An inter-observer variability study. *British journal of cancer*, 67(4), 773–775. DOI: 10.1038/bjc.1993.140.

Stefanidou, C. & C. Skordoulis (2014). Subjectivity and objectivity in science: An educational approach. *Advances in historical studies*, 3(04), 183. DOI: 10.4236/ahs.2014.34016.

Stelling, G.S. & S.A. Duinmeijer (2003). A staggered conservative scheme for every Froude number in rapidly varied shallow water flows. *International journal for numerical methods in fluids*, 43(12), 1329–1354. DOI: 10.1002/fld.537.

Supreme Court of New South Wales (2021). Queensland bulk water supply authority t/as seqwater v rodriguez & sons pty ltd [2021] nswca 206 (8 september 2021). https://classic.austlii.edu.au/au/cases/nsw/NSWCA/2021/206.html, Last accessed February 13 2024.

ter Horst, R., R. Alba, J. Vos, M. Rusca, J. Godinez-Madrigal, L.V. Babel, G.J. Veldwisch, J.P. Venot, B. Bonté, D.W. Walker & T. Krueger (2023). Making a case for power-sensitive water modelling: a literature review. *Hydrology and Earth System Sciences Discussions*, 2023, 1–31. DOI: 10.5194/hess-28-4157-2024.

Teuling, A.J., E.A. De Badts, F.A. Jansen, R. Fuchs, J. Buitink, A.J. Hoek van Dijke & S.M. Sterling (2019). Climate change, reforestation/afforestation, and urbanization impacts on evapotranspiration and streamflow in europe. *Hydrology and Earth System Sciences*, 23(9), 3631–3652. DOI: 10.5194/hess-23-3631-2019.

Thaler, T. (2021). Social justice in socio-hydrology—how we can integrate the two different perspectives. *Hydrological Sciences Journal*, 66(10), 1503–1512. DOI: 10.1080/02626667.2021.1950916.

Thieken, A., P. Bubeck & M.L. Zenker (2023). Fatal incidents during the flood of july 2021 in north rhine-westphalia, germany: what can be learnt for future flood risk management? *Journal of Coastal and Riverine Flood Risk*, 2. DOI: 10.59490/jcrfr.2023.0005.

Thompson, K.H. (1979). Margin of safety as a risk-management concept in environmental legislation. *Colum. J. Envtl. L.*, 6, 1.

Tsou, J.Y., A. Richardson & F. Padovani (2015). *Introduction: Objectivity in Science*, pages 1–15. Springer International Publishing, Cham. DOI: 10.1007/978-3-319-14349-1_1.

Turnhout, E., T. Metze, C. Wyborn, N. Klenk & E. Louder (2020). The politics of co-production: participation, power, and transformation. *Current opinion in environmental sustainability*, 42, 15–21. DOI: 10.1016/j.cosust.2019.11.009.

Uusitalo, L., A. Lehikoinen, I. Helle & K. Myrberg (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, 63, 24–31. DOI: 10.1016/j.envsoft.2014.09.017.

Pardo-del Val, M., C. Martínez-Fuentes, J.I. López-Sánchez & B. Minguela-Rata (2014). Franchising: the dilemma between standardisation and flexibility. *The Service industries journal*, 34(9-10), 828–842. DOI: 10.1080/02642069.2014.905925.

van den Berg, J. & M. van Lieshout (2022). De wraak van de Zuiderzee: renovatie Afsluitdijk kost honderden miljoenen extra door blunder Rijkswaterstaat. Published in the Volkskrant online on May 20 2022, https://www.volkskrant.nl/nieuws-achtergrond/de-wraak-van-de-zuiderzee-renovatie-afsluitdijk-kost-honderden-miljoenen-extra-door-blunder-rijkswaterstaat ba836a75/, Last accessed on July 27 2023.

van der Molen, D., F. Kortstee, J. Noort & H. Scholten (2002). Nen-norm voor modelleren in het

7

waterbeheer. *H2O: tijdschrift voor watervoorziening en afvalwaterbehandeling*, 35(16), 21–22.

van der Sluijs, J. P. (2002). A way out of the credibility crisis of models used in integrated environmental assessment. *Futures*, 34(2), 133–146. doi: 10.1016/S0016-3287(01)00051-9.

van Voorn, G., R. Verburg, E. M. Kunseler, J. Vader & P. H. Janssen (2016). A checklist for model credibility, salience, and legitimacy to improve information transfer in environmental policy assessments. *Environmental Modelling & Software*, 83, 224–236. doi: 10.1016/j.envsoft.2016.06.003.

van Waveren, R., S. Groot, H. Scholten, F. van Geer, H. Wüsten, R. Koeze & J. Noort (1999). *Vloeiend modelleren in het waterbeheer: Handboek Good Modelling Practice*. 99-05. Lelystad [etc.].

Venot, J. P., J. Vos, F. Molle, M. Zwarteveen, G. J. Veldwisch, M. Kuper, A. Mdee, M. Ertsen, R. Boelens, F. Cleaver, B. Lankford, L. Swatuk, J. Linton, L. M. Harris, J. Kemerink-Seyoum, M. Kooy & K. Schwarts (2022). A bridge over troubled waters. *Nature Sustainability*, 5(2), 92–92. doi: 10.1038/s41893-021-00835-y.

Venot, J. P., M. Zwarteveen, M. Kuper, H. Boesveld, L. Bossenbroek, S. V. D. Kooij, J. Wanvoeke, M. Benouniche, M. Errahj, C. D. Fraiture & S. Verma (2014). Beyond the promises of technology: A review of the discourses and actors who make drip irrigation. *Irrigation and drainage*, 63(2), 186–194. doi: 10.1002/ird.1839.

Verbeek, P. P. (2008). Morality in design: Design ethics and the morality of technological artifacts. In *Philosophy and design: From engineering to architecture*, pages 91–103. Springer. doi: 10.1007/978-1-4020-6591-0_7.

Verzijl, A., M. Vivek, A. Prayag, V. Srinivasan, C. Domínguez-Guzmán & M. Zwarteveen (2023). From divine to design: Unearthing groundwater practices in tamil nadu, india. *Water Alternatives*, 16(1), 153–170.

Voinov, A., K. Jenni, S. Gray, N. Kolagani, P. D. Glynn, P. Bommel, C. Prell, M. Zellner, M. Paolisso, R. Jordan, E. Sterling, L. Schmitt Olabisi, P. J. Giabbanelli, Z. Sun, C. Le Page, S. Elsawah, T. K. BenDor, K. Hubacek, B. K. Laursen, A. Jetter, L. Basco-Carrera, A. Singer, L. Young, J. Brunacini & A. Smajgl (2018). Tools and methods in participatory modeling: Selecting the right tool for the job. *Environmental Modelling & Software*, 109, 232–255. doi: 10.1016/j.envsoft.2018.08.028.

Vrontis, D. (2003). Integrating adaptation and standardisation in international marketing: the AdaptStand modelling process. *Journal of Marketing Management*, 19(3-4), 283–305. doi: 10.1080/0267257X.2003.9728212.

Vukovic, R. (2022). From conducting to inspiring: A qualitative research on servant leadership in the Corporate Service of Rijkswaterstaat.

Wamucii, C. N., P. R. Van Oel, A. Ligtenberg, J. M. Gathenya & A. J. Teuling (2021). Land-use and climate change effects on water yield from east african forested water towers. *Hydrology and Earth System Sciences Discussions*, 2021, 1–22. doi: 10.5194/hess-25-5641-2021.

Wang, H. H., G. Van Voorn, W. E. Grant, F. Zare, C. Giupponi, P. Steinmann, B. Müller, S. Elsawah, H. Van Delden, I. N. Athanasiadis, Z. Sun, W. Jager, J. C. Little & A. J. Jakeman (2023). Scale decisions and good practices in socio-environmental systems modelling: guidance and documentation during problem scoping and model formulation. *Socio-Environmental Systems Modelling*, 5. doi: 10.18174/sesmo.18563.

Ward, Z. B. (2021). On value-laden science. *Studies in History and Philosophy of Science Part A*, 85, 54–62. doi: 10.1016/j.shpsa.2020.09.006.

Watson, R. T. (2005). Turning science into policy: challenges and experiences from the science–policy interface. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1454), 471–477. DOI: 10.1098/rstb.2004.1601.

Watts, L. & D. Hodgson (2019). *Critical Social Science and Critical Theory*, pages 97–116. Springer Singapore, Singapore. DOI: 10.1007/978-981-13-3621-8_6.

Wears, R. L. (2015). Standardisation and its discontents. *Cognition, technology & work*, 17, 89–94. DOI: 10.1007/s10111-014-0299-6.

Werkowska, W., A. L. Márquez, R. Real & P. Acevedo (2017). A practical overview of transferability in species distribution modeling. *Environmental reviews*, 25(1), 127–133. DOI: 10.1139/er-2016-0045.

Wescoat Jr, J. L. & G. F. White (2003). *Water for life: Water management and environmental policy*. Cambridge University Press.

Wesselink, A., M. Kooy & J. Warner (2017). Socio-hydrology and hydrosocial analysis: toward dialogues across disciplines. *Wiley Interdisciplinary Reviews: Water*, 4(2), e1196. DOI: 10.1002/wat2.1196.

Wiarda, M., G. van de Kaa, N. Doorn & E. Yaghmaei (2022). Responsible innovation and de jure standardisation: an in-depth exploration of moral motives, barriers, and facilitators. *Science and Engineering Ethics*, 28(6), 65. DOI: 10.1007/s11948-022-00415-z.

Winkler, D., R. Hametner, T. Östreicher & S. Biffl (2010). A framework for automated testing of automation systems. In *2010 IEEE 15th Conference on Emerging Technologies Factory Automation (ETFA 2010)*, pages 1–4. DOI: 10.1109/ETFA.2010.5641264.

Winsberg, E. (2012). Values and uncertainties in the predictions of global climate models. *Kennedy Institute of Ethics Journal*, 22(2), 111–137. DOI: 10.1353/ken.2012.0008.

Wong, J., A. Fong, N. McVicar, S. Smith, J. Giambattista, D. Wells, C. Kolbeck, J. Giambattista, L. Gondara & A. Alexander (2020). Comparing deep learning-based auto-segmentation of organs at risk and clinical target volumes to expert inter-observer variability in radiotherapy planning. *Radiotherapy and Oncology*, 144, 152–158. DOI: 10.1016/j.radonc.2019.10.019.

Wood, A. W. (2001). The objectivity of value. *New Literary History*, 32(4), 859–881. DOI: 10.1353/nlh.2001.0058.

Woods, D. D. & N. B. Sarter (2000). Learning from automation surprises and going sour accidents. *Cognitive engineering in the aviation domain*, pages 327–353.

Xu, C., M. Rahman, D. Haase, Y. Wu, M. Su & S. Pauleit (2020). Surface runoff in urban areas: The role of residential cover and urban growth form. *Journal of Cleaner Production*, 262, 121421. DOI: 10.1016/j.jclepro.2020.121421.

Xu, L., P. Gober, H. S. Wheater & Y. Kajikawa (2018). Reframing socio-hydrological research to include a social science perspective. *Journal of hydrology*, 563, 76–83. DOI: 10.1016/j.jhydrol.2018.05.061.

Yang, A., D. Dutta, J. Vaze, S. Kim & G. Podger (2017). An integrated modelling framework for building a daily river system model for the murray–darling basin, australia. *International Journal of River Basin Management*, 15(3), 373–384. DOI: 10.1080/15715124.2017.1315814.

York, E., M. Brandt, S. N. Conley, M. Halpern, N. Mogul, E. Reddy, M. Stettler Kleine & D. Tomblin (SA).

Yusof, M. M., J. Kuljis, A. Papazafeiropoulou & L. K. Stergioulas (2008). An evaluation framework for health information systems: human, organization and technology-fit factors (hot-fit). *International journal of medical informatics*, 77(6), 386–398. DOI: 10.1016/j.ijmedinf.2007.08.011.

Zir, L. M., S. W. Miller, R. E. Dinsmore, J. Gilbert & J. Harthorne (1976). Interobserver variability in coronary angiography. *Circulation*, 53(4), 627–632. DOI: 10.1161/01.CIR.53.4.627.

7

Zurell, D., J. Franklin, C. König, P.J. Bouchet, C.F. Dormann, J. Elith, G. Fandos, X. Feng, G. Guillera-Arroita, A. Guisan, J. J. Lahoz-Monfort, P. J. Leitão, D. S. Park, A. Townsend Peterson, G. Rapacciuolo, D. R. Schmatz, B. Schröder, J. M. Serra-Diaz, W. Thuiller, K. L. Yates, N. E. Zimmermann & C. Merow (2020). A standard protocol for reporting species distribution models. *Ecography*, 43(9), 1261–1277. doi: 10.1111/ecog.04960.

Zwarteveen, M., M. Kuper, C. Olmos-Herrera, M. Dajani, J. Kemerink-Seyoum, C. Frances, L. Beckett, F. Lu, S. Kulkarni, H. Kulkarni, U. Aslekar, L. Börjeson, A. Verzijl, C. Dominguez Guzmán, M. T. Oré, I. Leonardelli, L. Bossenbroek, H. Ftouhi, T. Chitata, T. Hartani, A. Saidani, M. Johnson, A. Peterson, S. Bhat, S. Bhopal, Z. Kadiri, R. Deshmukh, D. Joshi, H. Komakech, K. Joseph, E. Mlimbila & C. de Bont (2021). Transformations to groundwater sustainability: from individuals and pumps to communities and aquifers. *Current Opinion in Environmental Sustainability*, 49, 88–97. doi: 10.1016/j.cosust.2021.03.004.

Zwarteveen, M. Z. & R. Boelens (2017). Defining, researching and struggling for water justice: some conceptual building blocks for research and action. In *Hydrosocial Territories and Water Equity*, pages 8–23. Routledge.

Data Statement

For the protection of privacy sensitive information, the interview data are unavailable for public release. The interview data are stored on the DANS Data Station Social Sciences and Humanities (Remmers, 2024, DOI: 10.17026/SS/AXTDZT).

Acknowledgements

This first paragraph is dedicated to the mistranscriptions I encountered (the italic parts are completely written by me) to give condolences to my own data. The data are from 2005 till the end of 2000 and were retrieved for a desktop study to understand the condition of different patients, since it helps to build a model to an immature level. This is done whenever we didn't have enough inflammation, because before we just reported on the terrible. This is of course a sibling project executed for the Bureau of Mythology and Piramidology. The basic plan was to have video guys look at the watermelons. But that's just lame, so they open the gates and fill an empty buffalo each year. The final conclusion: one person says 'Hi', when the other just said 'Bye'. *Thanks to these mistranscriptions, which I think form a hilarious and slightly incoherent story, I have laughed out loud and found one of the sources of humour to lighten my days.*

But in all seriousness, a PhD project is as crazy or as normal as the story I just concocted. To stay somewhat sane, I want to thank the people who've been my pillars, big or small. Because of these, I have been able to finish my PhD, to build on and grow and develop myself as an individual, as a professional and as a researcher. Even though I feel this list is quite extensive, I'm sure I've forgotten some people: sorry for that in advance!

First, thank you Lieke and Ryan for your supervision. Thank you for trusting me with this project and allowing me to make it my own and giving me the freedom to persue my own path, such as my research stay in Australia and my involvement in education. Lieke, I have appreciated your enthusiasm and passion for this topic and the fact that you got funding to do this research. We have had great discussions on the topic helping me further with my research. I have learned a lot due to your pragmatic, but attentive, way of working. Ryan, thank you for bringing a different perspective to the discussions. It was great to have someone challenge my own way of thinking. Also, I valued your scientific experience when writing and visualising my research.

Second, I thank all my interviewees for their time and input. Without all of you, my thesis would not exist.

Third, thank you Arjanne van Manen for illustrating my cover, the chapter images and the bookmarker. Thank you for taking my ideas and building on them and putting them on paper!

Fourth, I want to thank my co-authors, co-conveners, summer school buddies, the YHS committee, and other collaborators. A PhD is done on your own, but the research and outreach for it is not. Joseph, thank you for hosting and supervising me in Australia. And thank you to those in Australia who have showed me around or have given their perspective on the Australian system. Thank you to all my co-authors for providing input and giving feedback. Thank you to all who I've co-convened conference sessions with, I truly enjoyed it. YHS committee members thank you! Ruben, Linda, Fransje, Jerom, and Mel, it was great to set it up again together! And thank you to the current members for carrying it on.

Fifth, my colleagues. Within HWM, as a group, a friendly atmosphere has been created which fosters exchanging ideas and collaborating. For example, the tour de PhD was a brilliant trip embodying these two aspects. From HWM, I want to show my gratitude to some people specifically. First, the *vrouwenkamer*: Linda, Sjoukje, Rahel, Iris, Nazwa, Femke, Hadeel, Marjanne, Miranda, Tessa and the honorary members Joris and Bob. We have been able to share the highs and lows and lighten the mood when necessary. On top of that, it was so much fun to go to conferences with you! Of course, other colleagues have also enriched my PhD time with the laughter, coffee breaks and many different moments we've had.

Sixth, I want to acknowledge some groups and people outside of work who have listened and helped put my mind off things. I have met wonderful people in all these groups. So first, I might not have joined as often in the past few years (something about being abroad several times haha), but I love *kanokamp* and the *LKV*. It's such a supportive and attentive group. The lacrosse community has welcomed me with open arms, especially the different teams I've played for: in Wageningen, Nijmegen and now Bemmel and of course the *WDNT* A- and B-team. I appreciate how encouraging all of you have been and still are in my journey as a goalie, which started around the same time as my PhD. Also, Ivo Martinovic, we didn't speak a whole lot during my PhD, but whenever we did you always had a kind and uplifting word for me.

Seventh, I want to thank the people I've met through my volunteering work with *De Zonnebloem*. You gave me a confidence boost and allowed me to grow in a different environment, which I have been able to use during my PhD. At the local chapter in Wageningen, all the volunteers, especially *de keukenploeg* and the (former) board members, the people I visited, Hannie, Rose and mevr. van den Brink. Thank you for you warm welcome, your interest in my welfare and allowing me to clear my head and exchanging stories. Also, the multiple river cruises with *MPS De Zonnebloem*, those weeks were a joyful and true break for me and I know they will be in the future too. I thank the volunteers, the guests and the ship's crew for making these weeks so delightful and memorable.

Eighth, coincidental encounters have given me lasting friendships; dear friends who are as family to me. Therefore, thank you, Rosalie, Emese, Idske, Kikianne, Romy, Anna, Isabella and Ludovica. Thank you all for your support and careful listening to me and my concerns regarding anything. All of you have provided insightful advice and given me examples of how to tackle certain challenges. And just as or even more important, I have been able to relax and share beautiful experiences with you: from home-cooked dinners, to high teas, to spa days, to lengthy phone calls and to being welcomed in your (family) homes. I know I can always count on you.

Finally, my family. Even though my exact research topic may have been gibberish to you, all of you have always listened to me and supported me. *Dankjewel, papa, voor je liefde en trots die je gegeven hebt. Dat ik je kleine meid heb kunnen zijn voor even. Jouw liefde en trots zal ik altijd meedragen, zo ook nu ik deze mijlpaal heb bereikt.* Thank you, my darling brothers, *mijn kleine grote broer* and *mijn grote kleine broer*. I am so happy to share this moment with both of you as my paranymphs. I know I will always have your support as I have it today. Thank you both for the numerous chats, the practical advice, and indulging me in sharing and discussing my somewhat crazy thoughts. Esli, it has been a joy to have you join our family. Your caring, thoughtfulness and humour are wonderful traits. Thank you for your genuine interest and ability to listen. Last, but not least, mum, thank you from the bottom of my heart! Your unwavering support, the countless phone calls even when they were inconvenient for you (sorry...), your listening and questioning to let myself get to the answer, your encouragement, your letting us go. Thank you for being you and being such a brilliant example. I am so proud to call you all my family. I love you!

Statement of Authorship Contribution

JR = Janneke Remmers (WUR)

AT = Adriaan Teuling (WUR)
EN = Ehsan Nabavi (ANU)
JV = Jeroen Vos (WUR)
RD = Ruben Dahm (Deltares)
UP = Ulrike Proske (WUR)
AD = Arthur van Dam (Deltares)
JG = Joseph Guillaume (ANU)
LM = Lieke Melsen (WUR)
RH = Rozemarijn ter Horst (WUR)

Chapter 2

Conceptualisation:	LM, JR
Data collection:	JR
Data analysis:	JR
Data interpretation:	JR, in consultation with LM
Drafting of manuscript:	JR
Visualisation:	LM, JR, AT
Revision of manuscript:	LM, JR, AT

Chapter 3

Conceptualisation:	LM, JR
Data collection/assembling:	JR
Data analysis:	JR
Data interpretation:	JR, in consultation with LM
Drafting of manuscript, visualisation:	JR, in consultation with LM
Revision of manuscript:	AD, RD, LM, JR, AT

Chapter 4

Conceptualisation:	JG, LM, JR
Data collection/assembling:	JR, in consultation with JG
Data analysis:	JR
Data interpretation:	JR
Drafting of manuscript, visualisation:	JR
Revision of manuscript:	JG, LM, JR

Chapter 5

Conceptualisation:	JR, in consultation with LM
Drafting of manuscript:	JR
Visualisation:	UP, JR
Revision of manuscript:	RH, LM, EN, UP, JR, AT, JV

List of Publications

Peer-reviewed journal papers | In this thesis

Remmers, J.O.E., Teuling, A.J., Melsen, L.A. (2024). A modeller's fingerprint on hydrodynamic decision support modelling. *Environmental Modelling & Software.* DOI: 10.1016/j.envsoft.2024.106167

Remmers, J.O.E., Teuling, A.J., Dahm, R.J., van Dam, A., Melsen, L.A. (2024). Power to the programmer: Modeller's perspective on automating the setup of hydrodynamic models for Dutch water authorities. *Socio-Environmental Systems Modelling, 6, 18657*. DOI: 10.18174/sesmo.18657

Journal papers in preparation | In this thesis

Remmers, J.O.E., Guillaume J.H.A., Melsen, L.A. (2025). A True North: how does it evolve in hydrological modelling? A comparative study of modelling standardisation approaches in Australia and the Netherlands. *Environmental Modelling & Software (in review)*

Remmers, J.O.E., ter Horst, R., Nabavi, E., Proske, U., Teuling, A.J., Vos, J., Melsen, L.A. (2025). HESS Opinions: Reflecting and acting on the social aspects of modelling. *Hydrology and Earth System Sciences (submitted)*

Peer-reviewed journal papers | Other

Remmers, J.O.E., Teuling, A.J., Melsen, L.A. (2020). Can model structure families be inferred from model output? *Environmental Modelling & Software*. 10.1016/j.envsoft.2020.104817

van Hateren, T.C., Jongen, H.J., Al-Zawaidah, Beemster, J.G.W., Boekee, J., Bogerd, L., Gao, S., Kannen, C., van Meerveld, I., de Lange, S.I., Linke, F., Pinto, R.B., Remmers, J.O.E., Ruijsch, J., Rusli, S.R., van de Vijsel, R.C., Aerts, J.P.M., Agoungbome, S.M.D., Anys, M., Ramírez, S.B., van Emmerik, T., Gallitelli, L., Chiquito Gesualdo, G., Gonzalez Otero, W., Hanus, S., He, Z., Hoffmeister, S., Imhoff, R.O., Kerlin, T., Meshram, S.M., Meyer, J., Meyer Oliveira, A., Müller, A.C.T., Nijzink, R., Scheller, M., Schreyers, L., Sehgal, D., Tasseron, P.F., Teuling, A.J., Trevisson, M., Waldschläger, K., Walraven, B., Wannasin, C., Wienhöfer, J., Zander, M.J., Zhang, S., Zhou, J., Zomer, J.Y., Zwartendijk, B.W. (2023). Where should hydrology go? An early-career perspective on the next IAHS Scientific Decade: 2023–2032. *Hydrological Sciences Journal 68, 4, 529-541.* DOI: 10.1080/02626667.2023.2170754

Journal papers in preparation | Other

Pinto, R. B., Solitei, M., van der Ploeg, M., Remmers, J. O. E., Uijlenhoet, R., Akuoko-Gyimah, I., Van Emmerik, T. H. M. (2025). Mapping plastic waste entry points into the riverine environment

Global Environmental Change (submitted)

Publications in Dutch journals

Melsen, L., van der Bolt, F., Friocourt, Y., Hengeveld, G., van der Meijden, B., Remmers, J., Scholten, H., Talsma, M., Weerts, A., Wesselius, C., van der Zwet, J. (2022) Is goed modelleren goed genoeg? Kwaliteitsborging in modelleerstudies *Stromingen*

Curriculum Vitae

8 April 1996 **Born in Hengelo (OV), the Netherlands**

2013 - 2017 **BSc International Land and Water Management (free bachelor)**
Wageningen University, the Netherlands

2017 - 2020 **MSc Climate Studies** *(cum laude)*
Wageningen University, the Netherlands

2020 - 2024 **PhD Candidate**
Wageningen University, the Netherlands

Netherlands Research School for the
Socio-Economic and Natural Sciences of the Environment

D I P L O M A

for specialised PhD training

The Netherlands research school for the
Socio-Economic and Natural Sciences of the Environment
(SENSE) declares that

Janneke Oda Elisabeth Remmers

born on the 8th of April 1996 in Hengelo (OV), the Netherlands

has successfully fulfilled all requirements of the
educational PhD programme of SENSE.

Wageningen, 13 February 2025

SENSE coordinator PhD education

Dr Ir Peter Vermeulen

The SENSE Director

Dr Jampel Dell'Angelo

The SENSE Research School declares that **Janneke Remmers** has successfully fulfilled all requirements of the educational PhD programme of SENSE with a work load of 46.4 EC, including the following activities:

SENSE PhD Courses

- Environmental research in context (2020)
- Research in context activity: Organising Workshop called 'Model Etiquette' at RWS (2023)

Other PhD and Advanced MSc Courses

- Start to teach, Wageningen University (2021)
- Interview Techniques, Wageningen University (2021)
- Questionnaire Construction, Wageningen University (2021)
- The Choice: Un-box your PhD process and take charge of your performance, Wageningen University (2021)
- Grip op je Loopbaan: Young Career Professionals, Wageningen University (2022)
- Tour the PhD, Hydrology chair group (2022)

External training at a foreign research institute

- Process based Hydrological Modelling, University of Saskatchewan, Canada (2020)
- Ontario Branch Raven Hydrological Modelling training course, Canadian Water Resources Association, Canada (2021)
- Situating Hydrological Modelling, Humboldt University, Germany (2023)

Management and Didactic Skills Training

- Assisting in the tutorials of the BSc courses 'Water I' (2020-2024) and 'Water Quantity and Quality' (2020-2022)
- Teaching MSc course 'Catchment and Climate Hydrology' (2021-2024)
- Supervising two MSc students with thesis (2021-2022)
- Organising Writing Week for HWM chair group (2021-2022)
- Session organisation and convening at 6 conferences (3x EGU, 2x iEMSs, 1x MODSIM) (2021-2024)
- Secretary Young Hydrologic Society NL (2021-2024)

Selection of Oral Presentations

- *A Myriad of Motivations: how modelling decisions are made.* International Environmental Modelling and Software Society Conference 2022, 4-8 July 2022, Brussels, Belgium
- *A modeller's compass: how modellers navigate dozens of decisions.* EGU General Assembly 2023, 24-28 April 2023, Vienna, Austria
- *Building trust: How modellers navigate dozens of decisions*. MODSIM, 10-16 July 2023, Darwin Australia
- *A modeller's compass: how modellers navigate dozens of decisions.* EGU General Assembly 2024, 15-19 April 2024, Vienna, Austria

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Cover design: Janneke Remmers and Arjanne van Manen
Cover illustration: Arjanne van Manen
Layout design: Janneke Remmers, using $\LaTeX 2_\varepsilon$
Printed by: Luxor drukken & printen

Modellers as influencers?

Analysing practices and standards in hydrodynamic decision-support modelling

Janneke Oda Elisabeth Remmers