Universitat Autònoma de Barcelona

A value-based approach to agent-based simulation for policy assessment: an exploration in the water domain

Antonio Perelló Moragues

A value-based approach to agent-based simulation for policy assessment: an exploration in the water domain

Antonio Perelló Moragues

PhD in Computer Science

Line of Research: Artificial Intelligence

Advisor:
Pablo Noriega Blanco-Vigil

Tutor:
Sergi Robles Martínez

Universitat Autònoma de Barcelona

Artificial Intelligence Research Institute (IIIA-CSIC)
Barcelona, 2020

Amb el suport del Pla de Doctorats Industrials de la Secretaria d’Universitats i Recerca del Departament d’Empresa i Coneixement de la Generalitat de Catalunya (2016 DI 043).

En aquest projecte de doctorat industrial hi han col·laborat Aqualia, l’IIIA-CSIC, i la Universitat Autònoma de Barcelona.

Abstract

Policy-making is a value-driven activity. From the assessment of states of the world to the choice of the appropriate actions to do under given circumstances, values are present in deciding whether something is good or not. Thus, values are involved in policy design, but also in the behaviour of those individuals who are going to be affected by a policy. Since these value-driven choices in policy-making may have significant social consequences, it is advisory to assess policies prior to their enactment. Assuming that agent-based simulation is a powerful methodology for this purpose, the need of a conceptual framework that includes values into policy simulator systems is clear. Interestingly, this leads to explore the role of values in problem design from the ground, especially the model and the outcome assessment. In order to imbue values into computational models, a methodology is proposed: contextualise and translate abstract values into state indicators, and build the simulator around these representations (from the set of interventions to the their assessment). This methodology proposal is illustrated in the water domain, and is used to address two case studies by means of simulation: the first, the *modernisation* of farmer communities in Spain; and second, the interaction of policy actors that lead to management policy shifts in the urban context. Finally, the contributions of the framework in issues that are raising concerns in Artificial Intelligence nowadays, as well as potential contributions to water management, are explained as future work lines.

Keywords: agent-based simulation · policy-making · policy assessment · values · computational social sciences · social simulation · socio-hydrology

Resum

La creació de polítiques públiques és una activitat motivada per valors. Des de la avaluació d'estats del món a l'elecció de l'acció apropiada sota unes circumstàncies determinades, els valors estan presents a l'hora de decidir si quelcom és bo o no. Així, els valors juguen un paper en el disseny de polítiques, però també ho fan en el comportament dels individus que seran afectats per una política. Donat que aquestes decisions motivades por valores poden tenir conseqüències socials significatives, és recomanable avaluar les polítiques públiques abans de la seva promulgació. Assumint que la simulació basada en agents és una metodologia potent per aquest propòsit, la necessitat d'un marc conceptual que inclogui valors en els sistemes de simulació de polítiques és clara. De manera interessant, això porta a explorar el rol dels valors en el disseny de problemes des dels fonaments, especialment pel que fa als models i a l'avaluació de resultats. Amb la finalitat d'infondre valors en els models computacionals, es proposa una metodologia: contextualitzar i traduir els valors abstractes en indicadors d'estats, i construir els simuladors envoltant d'aquestes representacions (des del conjunt d'intervencions a la seva avaluació). Aquesta proposta metodològica és il·lustrada en el domini de l'aigua, i llavors és utilitzada per abordar dos casos d'estudi mitjançant simulació: en primer lloc, la modernització de comunitats de regants a Espanya; en segon lloc, la interacció d'actors polítiques que condueix a girs en les polítiques de gestió a l'àmbit urbà. Finalment, les contribucions del marc conceptual en qüestions que estan suscitant preocupacions actualment en l'Intel·ligència Artificial, així com contribucions potencials a la gestió de l'aigua, s'expliquen com línies de treball futures.

Keywords: simulació basada en agents · creació de polítiques públiques · avaluació de polítiques públiques · valors · ciències socials computacionals · simulació social · socio-hidrologia

Resumen

La creación de políticas públicas es una actividad motivada por valores. Desde la evaluación de estados del mundo a la elección de las acciones apropiadas bajo unas circunstancias determinadas, los valores están presentes a la hora de decidir si algo es bueno o no. Así, los valores juegan un papel en el diseño de políticas, pero también lo hacen en el comportamiento de los individuos que van a ser afectados por una política. Ya que estas decisiones motivadas por valores pueden tener consecuencias sociales significativas, es recomendable evaluar las políticas públicas antes de su promulgación. Asumiendo que la simulación basada en agentes es una metodología potente para este propósito, la necesidad de un marco conceptual que incluya valores en los sistemas de simulación de políticas es clara. De modo interesante, esto lleva a explorar el rol de los valores en el diseño de problemas desde sus cimientos, especialmente en lo que atañe a los modelos y a la evaluación de resultados. Con el fin de imbuir valores en los modelos computacionales, se propone una metodología: contextualizar y traducir valores abstractos en indicadores de estado, y construir los simuladores alrededor de estas representaciones (desde el conjunto de intervenciones a su evaluación). Esta propuesta metodológica es ilustrada en el dominio del agua, y es usada luego para abordar dos casos de estudio mediante simulación: en primer lugar, la modernización de comunidades de regantes en España; en segundo lugar, la interacción de actores políticos que conlleva a giros en las políticas de gestión en el ámbito urbano. Finalmente, las contribuciones del marco conceptual en cuestiones que están suscitando preocupaciones actualmente en la Inteligencia Artificial, así como contribuciones potenciales a la gestión del agua, se explican como líneas de trabajo futuras.

Keywords: simulación basada en agentes · creación de políticas públicas · evaluación de políticas públicas · valores · ciencias sociales computacionales · simulación social · socio-hidrología

List of publications and communications

- Perello-Moragues, A., Noriega, P.: A Playground for the Value Alignment Problem. In: Martínez-Villaseñor, L., Batyrshin, I., Marín-Hernandez, A., Ponce, H. (eds.) Advances in Soft Computing. MICAI 2019. Lecture Notes in Computer Science, vol 11835. pp. 414–429. Springer International Publishing, Cham (2019)
- Perello-Moragues, A., Noriega, P.: Using Agent-Based Simulation to understand the role of values in policy-making. In: Verhagen, H., Borit, M., Bravo, G., Wijermans, N. (eds.) Advances in social simulation — Looking in the mirror. Springer (In Press)
- Perello-Moragues, A., Noriega, P., Padget, J., Verhagen, H.: Value-driven policy-making as a socio-cognitive technical system. In: First international workshop on socio-cognitive systems (SCS18) in the International Conference on Autonomous Agents and Multiagent Systems (AAMAS18) (Communication, 2018)
- Perello-Moragues, A., Noriega, P., Poch, M.: Modelling contingent technology adoption in farming irrigation communities. Journal of Artificial Societies and Social Simulation 22(4), 1 (2019)
- Perello-Moragues, A., Noriega, P., Popartan, A., Poch, M.: Modelling policy shift advocacy. In: Proceedings of the Multi-Agent-Based Simulation Workshop (MABS) in the International Conference on Autonomous Agents and Multiagent Systems (AAMAS19) (In Press)
- Perello-Moragues, A., Noriega, P., Popartan, A., Poch, M.: On three ethical aspects involved in using agent-based social simulation for policy-making. In: Proceedings of the Social Simulation Conference 2019 (In Press)
- Sierra, C., Osman, N., Noriega, P., Sabater-Mir, J., Perello-Moragues, A.: Value alignment: A formal approach. In: Responsible Artificial Intelligence Agents Workshop (RAIA) in the International Conference on Autonomous Agents and Multiagent Systems (AAMAS19) (Communication, 2019)
- Teze, J.C., Perello-Moragues, A., Godo, L., Noriega, P.: Practical reasoning using values: an argumentative approach based on a hierarchy of values. Annals of Mathematics and Artificial Intelligence 87(3), 293–319 (2019)

Acknowledgements

None of this would have been possible without the help of many people, and I wish to thank them all.

To Pablo Noriega, for advising this thesis. To, Manel Poch, who has been virtually a co-adviser of this dissertation.

To Alexandra Popartan, for sharing the adventures and misadventures of this project.

To all the people in the IIIA. To those people who I collaborated with at some point: Lluis Godo, Jordi Sabater, Carles Sierra, Nardine Osman and Juan Carlos Teze. Especially, to the other students in the IIIA. To those who I could only spent a short time with: Xavi, Albert, Sergi, Ewa, Kemo and Paula. To those who I spent the most time with: David, Ramón, Adán, and Oguz.

To the people in Aqualia: Pedro Rodríguez, Eva Martínez, Frank Rogalla, Pilar Icaran, Antonio Giménez, Gracia Martínez, Beatriz Del Castillo, and Salvador Salse. I appreciate the freedom they gave me to research, their interest, and the support of the technical staff when we asked for it.

To Sergi Robles, who was there to guide me in the administrative maze.

To Harko Verhagen and Julian Padget. Not only for their suggestions in my work, but also for sharing those dinners in places I would have never gone.

To Francisco Alcón, without whom I would not have published my first paper.

And last, but most important, to my family: A na Xisca per ser devora jo. A mon pare, a na Marga, a en Tòfol i a n'Isa, a en Pere i a n'Andreu, a en Pau i a n'Aina, a na Carley i a sa família.

– Toni

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Motivation and Overview

> *"A truly smart system would find a way to turn us into more reflective, caring, and humane creatures"*
> – Evgeny Morozov

It is well-known that the point of public policies is to achieve a *better* state of affairs. Public policies are, in plain terms, plans of action that address collective problems [65] by means of coordinated social action. However, how social coordination can be effectively achieved? In any society or organisation, stakeholders may have different values and motivations, which can be a source of disagreement, conflict, and misdirection. Such misalignment may happen at different levels. At the level of interactions, stakeholders who are coordinated by the policy may pursue different goals and have different skills and personalities. At the policy-maker level, the policy may be designed in a way that is not compatible with the goals of other stakeholders, or even may have indirect unintended effects.

For instance, rural socio-economic development has been a recurrent topic of Spanish public policies which has been traditionally linked to irrigation agriculture. The national hydraulic policy of the past twentieth century was based on increasing the supply of water resources and the transformation of irrigation, in order to generate employment and curb rural depopulation [148]. Nowadays, it has focused on promoting the collective adoption of modern irrigation systems (i.e. pressurised networks and sprinklers or drip pipes) of farming communities by means of subsidies in order to increase agricultural productivity. However, although this intervention has worked in some cases, in others it has been completely rejected invoking tradition, self-management or even aesthetics [215]. In these cases, the ethical reasoning behind the public policy may not match with the goals and morality of the target groups, making the policy "uneffective" to achieve its desired change. But it can also happen that, despite successfully appealing to the target groups, the attempted solution leads to negative unintended consequences. The paradigmatic example is the case of India under the British rule. The colonial government were concerned about the number of venomous snakes and decided to offer an economic reward for every dead snake. Although it initially worked, the problem did get worse. The reason was that people did breed snakes at home in order to earn an income. When the government realised that behaviour and terminated the policy, caused the snake breeders to set the animals free, which made the problem worse that was at the beginning.

These stories illustrate some potential failures of public policies. But they also show that public policies are based on models that are built on value-driven considerations (from the policy ends to the policy means), although these values are not always easily identified, or made explicit.

The core of this dissertation is the acknowledgement that public policies may have significant social consequences and are based on value-driven choices. Its contents are an attempt to make that insight operational by devising a framework to represent policies, exploring its applicability for policy assessment and identifying opportunities for further work.

A second premise behind this dissertation is that the understanding of societal problems can be enhanced through computer systems based on simulation (see [173]), and that agent-based simulation, in particular, is a convenient approach to support policy-making and conduct *ex ante* policy assessment [91].

The challenge is that some fundamental questions about values and value-driven behaviour —which are essential for simulation-based policy assessment— have not been systematically addressed in agent-based simulation. For example: How can one assess the *value* of a particular social outcome or state of affairs? Why are some policy means (an incentive, a promotional campaign, new regulations) more effective towards policy goals? How to identify the relevant side effects of a policy? What are the features of the target groups of a given policy proposal that relate to the satisfaction with respect to a *value*? What are the effects of the policy on those values that are not considered *relevant*?

These are topics of critical importance for policy-making and, therefore, for policy simulator systems, because it is known that assumptions and drivers behind policy decisions are based not only on factual knowledge but on ethical beliefs and intuitions as well [237, 134, 247]. This motivation leads to consider the notion of values, a construct that underlies many of the human decisions [221, 110]. Taking values into account may contribute to improve the design of simulator systems for decision-support: (i) by enabling a more meaningful use of these tools (affording the user to consider its own values); (ii) by enabling a more expressive modelling of the motivations of the agents; and (iii) by revealing the social effects of policies with regard to diverse social values (indicating the social significance of policy decisions and reinforcing transparency and accountability).

Thus, this dissertation proposes a conceptual framework that makes the notion of value preeminent by providing, first, the key constructs to specify **value-driven policy simulator systems** and, second, support methodological guidelines to model actual policies. A significant outcome of this proposal is that such value-driven policy simulators constitute a convenient sandbox for the “value-alignment problem” [211] (see Chapter 11).

1.1 Values and policy simulator systems

1.1.1 Values and social outcome

It is common that, when mentioning the *value* of a state of affairs, it induces to think about wealth-related concepts. This is not surprising, since the understanding of economic progress in western democracies has been historically reduced to GDP-based economic growth, with many political and academic figures contributing to establish this particular view.

However, and this is key, *value* can be generated from a broader system of values, thus revealing the range of diverse aspects that are present in the social welfare. Paraphrasing Mazzucato [161], if *value* is not questioned in public debates, only some activities and public services will be deemed to be value-creating, while others will not. And it is obvious that these understandings will guide policy-makers’ strategic decisions and governance settings.

Understanding how values define societal goals, and having a proper way to implement and assess values in simulator systems, may help to have a better view about how contributing to social well-being and create proper stories for that matter. Moreover, it can further contribute to transparency and trustworthiness of decision-makers, which is indeed a competitive advantage. This dissertation suggests a methodology to contextualise values in a domain of interest and then translate them into indicators that support their assessment (see Chapter 8).

1.1.2 Values and stakeholders

Besides supporting the evaluation of states of affairs, values are also involved in decision-making, as they are related to motivation and goals [183, 2]. In practice, the behaviour of stakeholders is deeply related with the values they hold.

Values are present both at the individual level (i.e. guiding some cognitive functions and, eventually, behaviour), but also at the social level (i.e. constituting beliefs about what the society should ensure and preserve) [207, 226, 270]. This is a relevant interplay when addressing collective problems, since values are involved in policy-making (i.e. collective coordination) and in individual behaviour —especially, in acceptance of policies and states of affairs (see [269, 44]).[1]

In this dissertation, it is assumed that values shape the cognitive functions that agents perform within a particular context and domain, namely by: (a) guiding perception of the environment (i.e. defining what is relevant and deserves attention); (b) guiding the evaluation of objects (especially, the assessment of (objective) states of affairs and actions' outcomes); and (c) guiding individuals' decision-making (i.e. what goals are set and what are the best courses of action) (see the examples in Chapter 8).

1.1.3 Values and policy models

Human beings use models to understand and interact with a complex environment [126, 62, 30] and provide explanations and inferences about diverse phenomena [36]. Given that policy-making has to address collective problems in complex societal systems, decision-makers may use policy models to convert a messy problematic situation into a structured well-defined problem that affords to design interventions [115]. Simulator systems may further support decision-making, allowing more coherent and informed decisions, because they are able to reproduce characteristics of complex systems satisfactorily (e.g. nonlinear dynamics, network interaction, emergence) and generate new insights [173].

However, policy models, since they simplify the world, may be incomplete and biased, which leads to ethical concerns on their design and use [19]. Knowing that values shape the perception and assessment of the environment [111, 183], their role in model-building for policy simulator systems is critical and therefore must be further explored (Chapter 3).

1.2 A challenging problem for Artificial Intelligence

Although Artificial Intelligence has focused historically on paradigmatic aspects of individual intelligence (e.g. learning, reasoning, image recognition), the field of social intelligence have gradually become mainstream, making coordination and governance salient topics (see [52, 59, 68, 53, 18, 12]). In a sense, public policies achieve a particular state of affairs by means of a diversity of governance mechanisms (e.g. incentives, regulations, rhetorical messages, etc.). In this context agents' values are relevant because they are involved in how these agents behave and react to these instruments.

This approach supports the possibilities of value-imbued AI-backed systems that may be useful for: (i) building virtual testbeds for technology and policy assessment; (ii) designing digital platforms for participatory policy-making or policy monitoring; and (iii) computational approaches to solve "policy problems". Another area where these ideas have been incorporated is in value-based design methodologies (e.g. [196]). However, because of the concerns associated with the autonomy of artificial entities, the notion of value as a source for control becomes evident. Thus imbuing moral values in the

---

[1] While writing this thesis, the Joint Research Centre of the European Commission is working on an initiative called "Enlightenment 2.0", whose research agenda includes science of values in the political process, in order to cover these topics (`https://ec.europa.eu/jrc/en/enlightenment/call-experts-science-of-values`).

decision models of autonomous agents would foster self-governance and provide grounds for projecting trust on those agents. Likewise, imbuing values in the mechanisms that govern a multiagent system (as values are imbued in a policy) provides grounds for assessing the trustworthiness of that system. These are the motivations behind many ethical AI initiatives (see [121]) and the above mentioned value alignment problem (Chapter 11). In this dissertation, a methodology to imbue and assess values in computational models is proposed and illustrated (Chapters 4 and 8).

1.3 Focus on the water domain

Water is the most essential resource on Earth. It is indispensable for life and it is present in many facets and activities of human civilisations (i.e. agriculture, public health, energy generation, culture and religion, etc.). Hence, water is at the heart of several social and ecological conflicts along the history of humankind. Water is a resource with multiple uses and stakes, whose management impacts across the interwoven socio-economic and ecological systems.

This takes special importance on water services, that are between the most fundamental public services in human settlements (which include the supply of drinking water and the sanitation of wastewater). Citizens' well-being is directly linked to the governance and management of public water services, as they use water for diverse basic needs, such as hygiene and sanitation.

Water governance and management issues are particularly urgent nowadays, since global conditions are changing rapidly (for instance, global warming caused by climate change or population growth) and may pose a severe threat to humanity [48, 204], especially in the water domain [254, 232]. In fact, there has been already phenomena with harmful effects for the social welfare, as seen in Cape Town or Australia (see [51, 266, 241]). However, many attempted solutions to these threats rely on the enactment of instruments that can lead to unintended effects on some social values, leading to social acceptance issues (see [269, 44]). For instance, water rationing, with the goal of preserving a minimum level for critical services and the environment, may not be socially sustainable if not appropriately implemented.

For this reason, addressing such urgent water-related threats properly depends on enhancing the understanding of collective action and social coordination to support decision-making in public policies. In this spirit, a new field called socio-hydrology [239] aims at studying the water cycle together with social behaviour, for which AI-approaches may be extremely meaningful. Public policies must preserve and improve the welfare of population in challenging situations that involve value dilemmas and critical consequences.

Given this purpose, some research lines in water policy are increasingly focusing on values and preferences of citizens to improve the responsiveness to their needs and ethical standards (see [41, 152]. Values are broadly studied in environmental habits adoption [67, 209], but they are also relevant to determine the other motivations of individuals [221, 183]. Thus, some scholars point out that much more knowledge in households' "soft" characteristics with regard to water is needed in order to build better models (e.g. values, lifestyles, preferences, motivations, etc.) [41, 127].

Building on these insights, water is known to be critically involved in social welfare, making it a relevant domain where to explore policy-making and the role of values. In this spirit, this dissertation explore how to include these aspects in agent-based models for water policy-making. In particular, it illustrates how to use values into water models (Chapter 8), and examines two case studies by means of simulation: the *modernisation* of farmer communities (Chapter 9) and the shifts in urban water policies (Chapter 10).

1.4 Dissertation

1.4.1 General objective and approach

This dissertation is based on two main premises: (i) that policy-making is a value-driven social process; and (ii) that agent-based simulation is an appropriate methodology to support policy assessment. The objective of this thesis is to outline a conceptual framework for policy simulator systems so that these take values into account. This framework emphasises the main elements of a "policy problem" (namely, the *domain* of intervention, the *stakeholders*, the *values*, and the intervention itself or *policy-schema*).

This thesis takes a classical Artificial Intelligence approach: that is, it focuses on the development of a conceptual framework that can be used to model "policy problems" in computer systems in order to provide tools to human experts. Paraphrasing Simon [236], and Russell and Norvig [212], Artificial Intelligence consists of *intelligence* and an *artifact*; in other words, a model to represent and address a problem or task, and the computer that implements it and performs such function. As a consequence, the conceptual framework aims at providing a high-abstraction language that facilitates modelling such problems in computer systems.

In more precise terms, the dissertation addresses the following questions:

Q1. How value-driven social coordination problems (i.e. "policy problems") can be represented; ;
Q2. How values can be used and measured in computational models for simulation;
Q3. How social outcomes can be assessed with respect to multiple values through agent-based simulation;
Q4. How to assess a policy through agent-based simulation;
Q5. How agent-based modelling can engage in sensitive design (like *value-sensitive design* [196]);
Q6. Explore these topics for the water domain.

1.4.2 Structure

This dissertation organises the aforementioned matters in four parts:

- Part I consists of this introduction and Chapter 2, that reviews literature on (i) values in individual behaviour (i.e. motivations, decision-making, political behaviour, etc.) and in public policies (i.e. social standards, strategic decision-making, political and ethical concerns, etc.); on (ii) policy assessment procedures (especially, the design of policy models and the analysis of the expected social effects); and (iii) the use of agent-based simulation to support the policy assessment and inform policy-making.

- Part II develops the conceptual framework. In particular, Chapter 3 explains the assumptions on which the framework is based (i.e. public policies, values, and modelling for simulation). Then, Chapter 4 makes those intuitions more precise by introducing formal definitions of the core concepts involved in value-driven policy simulator systems. Afterwards, Chapter 5 expands the conceptual framework to include new affordances for agents so as to model policy shifts. Finally, Chapter 6 discusses methodological guidelines to address the ethical concerns that are associated with the design and use of value-driven simulation.

- Part III deals with the use of the conceptual framework in a particular policy domain. Chapter 7 introduces water as policy domain to illustrate the contents of Part II, including a brief literature review on values and simulation in the water domain. Chapter 8 illustrates how the conceptual framework is used and, specifically, how to make abstract values operational for computational models in the water domain. The other two chapters in this part are two case studies in which simulation is used to support decision-making. First, Chapter 9 presents a case study about the modernisation of farming communities in Spain. In this case study, the model is built on actual data from the modernisation process in two communities and show how a single value (economic profit) explains satisfactorily the adoption process in these communities. It also illustrates how side-effects on other values may be assessed, and explains the need of additional values to account for the adoption processes of communities where other circumstances prevailed. Second, Chapter 10 is used to illustrate other features of the framework in the context of urban uses of water. The simulator models how opinion-makers advocate for policy shifts and the alignment of their values with those of the households determine the support given to shift proposals. In this case the model includes different stakeholder roles (namely, households and opinion-makers) with different capabilities and with profiles that include several values. Simulations explore the interplay between ideological affinities of households and opinion-makers with respect to policy instruments and demographic composition.

— Finally, Part IV gives a perspective on future work. Chapter 11 points at extensions of the results, namely, the contribution of the conceptual framework for aligning AI-backed systems with values. It describes the main contents of two papers whose formalism is based on the conceptual framework (namely, a value-based approach for argumentation [252] and a formal foundation for value-driven normative multiagent systems [234]). In Chapter 12, some potential applications of the conceptual framework in water management and environmental policy-making are proposed. Lastly, Chapter 13 presents the main conclusions of this dissertation.

1.4.3 Publications and communications

The publications and communications elaborated during this thesis are reflected on the contents of the chapters as follows:

— The works from [188], [187] and [186] are preliminary works on the assumptions (Chapter 3) and concepts of the framework (Chapter 4). In [188] and [187] the main elements of the framework were outlined (e.g. stakeholders, values, domain, policy schema), which were further developed and formalised in [186]. This last publication, moreover, suggested the potential use for exploring values in AI-backed systems (Chapter 11). This dissertation reorganises in a coherent manner the contents of those publications.

— The case study in Chapter 9 was published in [189]. Although in that publication the elements of the framework of this dissertation were not strongly emphasised, the process of modelling the case study for simulation did follow its guidelines (which provided insights to structure the framework, as well).

— The work in [190] served to extend the framework by exploring different stakeholder roles and capabilities in order to address case studies where values and politics are a main component of the system, which drives agent interaction and eventually may produce policy shifts. This theoretical exploration has been compiled in Chapter 5. Besides, this work also presented a preliminary example of the applications in mind while developing such framework extension, which has been the basis for the case study in Chapter 10.

— In [191] the idea was to discuss how values are instantiated in a domain and how different types of values are taken into account in the process from modelling a policy to using a simulator system, which is reflected in Chapter 8 and Chapter 6 respectively.

- Finally, the two collaborations are formalisms based on the conceptual framework and are included in Chapter 11: the work from [252] that describes a value-based approach for argumentation; and the work from [234] provides a foundation for value-driven normative multiagent systems.

Chapter 2

Background

Before presenting the conceptual framework for policy simulator systems, it is convenient to examine the literature that supports the major foundations of the work. This includes background on three main pillars:

- The notion of values: what they are and how they are present in human cognition and human behaviour. In another view, how they assist in public policies and strategic decision-making.
- The process of policy assessment: what steps are to be taken in order to assess a policy designed to address a collective problem.
- The use of agent-based simulation for policy-making: what it consists of and why it is a promising methodology for policy assessment.

2.1 Values

Values are transcendental beliefs that refer to desirable goals and motivate action [225]. Generally speaking, values are one's principles that define what is important in life and guide one's judgement of what is right and what is wrong, or whether something is good or bad. Values are often viewed as one's ethical beliefs that determine standards that guide the conduct of the person. In a more practical view, values are involved in assessing the desirability of states of affairs and outcomes and deciding whether one action is preferable to another, eventually guiding the behaviour of individuals.

Taking this into account, values are not only relevant for the behaviour of individuals, but also at a social scale. The outcome of the social behaviour (that is, the behaviour of the population of individuals as a whole) is also influenced by values, from two different perspectives. On the one hand, individuals make decisions, interact between each other, and may also participate in political affairs. On the other hand, managers, public servants and policy-makers invoke values when making decisions in public affairs in behalf of the society. Hence, values are present in shaping the social outcome —and, for this reason, it is convenient to express and distinguish values in a practical sense.

2.1.1 Values as a multidisciplinary construct

Traditionally, values have been present in normative disciplines like ethics and politics, when asking questions about what constitutes good social outcomes and how societies should be governed for that purpose. Values are also present in many other disciplines, especially for analytical purposes. They have been covered in psychology [221, 183, 111], engineering [195, 196], economics [161], and sociology [108], to mention some of the major fields today.

This has lead to multiple perspectives and different definitions for values. As a consequence, the definitional inconsistency and the multiple methodological approaches across academic fields constitute one of the main problems for the science of values [206, 110].

It is broadly accepted that values are cognitive constructs. They cannot be proved wrong, they cannot be measured directly, and they express inner principles of what individuals consider good and right. That is why values have been historically related to ethics and moral reasoning [108, 224], and then to studying the psychological concepts with regard to the conduct of individuals [183]. However, although values are cognitive constructs, they are useful for modelling socio-cognitive systems: they provide a description of an aspect of human cognition that is extremely useful for modelling social behaviour (which includes also an ethical reasoning of individuals) and for setting standards for policy assessment —useful for the design of simulator systems for that matter.

In addition, values are also relevant from a social perspective. Since values define what is right and good for individuals, they will eventually determine their behaviour in society and their concern for others [99]. This takes special relevance when considering agents that act in social spaces where other beings and entities coexist: for instance, from a socio-cultural perspective, values can determine the cooperation and interaction of individuals (e.g. altruism, reciprocity, etc.).

For the sake of practicality, values are used as standards to define the *value* of some object in some disciplines. In this spirit, *value* —in singular— is the inherent degree of worth, importance or desirability assigned to an object (e.g. event, action, or outcome), which is elicited from a system of *values* —in plural. Accordingly, *assigned value* refers to the contribution of an object to meeting a specific goal or objective [60]. In some cases, this view can help not only to identify the *value* (i.e. gain) of something, but also the *cost* (i.e. loss). However, this approach is subject to severe criticism, provided that the system of *values* trough which the *value* is elicited is overshadowed —thus pushing aside the debate of what *values* should be considered when assessing something.

2.1.2 Values from an agent perspective

When speaking about agents' cognition, the notion of values emphasise some particular aspects present in behaviour.

Many sources in psychology concur that values play an active role in the intentional human behaviour —regardless of whether reasoning about them is conscious or not—, as they play an important role in setting *goals* and elicit *preferences* [183]. It is said that human beings are goal-directed, understanding goals as desired outcomes [2]. Individuals perceive the environment and its discrepancies with their goals, which may motivate them to pursue action. In other words, individuals assess the state of affairs with regard to their values, the outcome of which may motivate them to take some particular course of action. For this reason, there is abundant support on that values are connected with motivation [207, 222, 183], which relates to what individuals pursue and how they pursue it in terms of intensity and persistence [183]. Alternatively, values are often approached as internalised norms (such as habits) that determine agents' courses of action (that is, influencing their behaviour not due to a deliberative analysis of an assessment, but rather due to following an internal rule) [209].

Although goal-setting is commonly considered to be a conscious and intentional process, goals can be established, or at least influenced, by external elements. Decision situations can be framed in different ways that give rise to different preferences [130]. This is particularly important in social systems, as values play a great role in framing social issues and collective problems [143, 144].

It is convenient to say that values are not the only cognitive constructs involved in eliciting goals and preferences. Values are part of the cognitive system and are intrinsically connected with other cognitive components like mental models —that are used to understand and interact with the environment [126] and provide explanations and inferences about diverse phenomena [36]—, personal traits [183], emotions [223, 25], and attitudes [135, 110], besides being affected by diverse cognitive biases [130].

One of the significant characteristics of values for modelling-purposes is that they are largely stable internal cognitive constructs. It has been suggested that values arise from lived experiences and upbringing, and they only change under singular events [183].

Cognitive view of values: perception, evaluation, and decision-making

In plain terms, values are considered to be constructs that guide behaviour by acting as individual-level cognitive heuristics [270], standards of preference [207] and moral intuitions [110]. Especially, they are related to motivation, thus articulating reasons to initiate and perform intentional behaviour in order to satisfy the (human) basic needs of the individual [203, 183, 225].

Schwartz and Bilsky [225, 219] provided the most exhaustive and mainstream definition of values: (1) values are beliefs (2) that refer to desirable goals that motivate action, (3) transcend specific situations, (4) guide selection and evaluation of actions, policies, people, and events, (5) are ordered by relative importance, and (6) whose relative importance guides action.

In practice, values serve to make value judgements or assessments —thus, they are, in essence, an evaluative construct. Individuals decide which course of action is preferable by using values —as they inform about what constitute appropriate behaviours and desirable states of affairs. Naturally, many situations involve a choice between *a priori* equally-right alternatives, presenting an unclear picture of what course of action is preferable to take, which is, in essence, an ethical dilemma.

However, individuals do not reason explicitly on values in general terms. It has been suggested that only situations that are new or complex enough lead to an active value-based reasoning, while routine is based on habits —which would have been adopted by an active value-based reasoning process in the first place [183].

Given a finite set of options in response to a challenge or opportunity in the environment, individuals bring their potential outcomes to mind and compute their expected *value*, in order to determine whether these outcomes should be set as goals [2]. Two additional constructs are relevant in this process: (i) the rewarding *value* of the outcome and (ii) the *expectation* of being able to attain it. Yet, individuals are incapable of exploring every course of action and its consequences [235], besides being their judgements often biased and based on heuristics [130]. As it is well-known, human rationality is *bounded* and affected by emotional and intuitive processes (in which, values would play a relevant role).

In this dissertation, it is assumed that values shape the cognitive functions that agents perform within a particular context, namely by: (a) guiding perception of the environment (i.e. defining what is relevant and deserves attention); (b) guiding the evaluation of objects (especially, the assessment of (objective) states of affairs and actions' outcomes); and (c) guiding individuals' decision-making (i.e. what goals are set and what are the best courses of action).

Theoretical approaches and taxonomies of individual values[1]

Several approaches have focused on classification of values from a cognitive perspective of the individual. Although there is no consensus on taxonomies, some works stand out.

---

[1] Most of the following models are based on surveys to empirically elicit individual values (see [104]).

One of the most known motivational theories is the one proposed by Maslow [159]. Although this theory does not mention values explicitly, it resulted in much work that related them to motivation and human needs (see [203]). Maslow's theory defines a hierarchy of human needs that is represented as a pyramid: needs lower down must be satisfied before individuals can attend to the needs higher up the pyramid (Fig. 2.1). According to the theory, there are five hierarchical levels: (i) physiological; (ii) safety; (iii) love and belonging; (iv) esteem; and (v) self-actualisation. Later, Maslow thought on adding a sixth level to the pyramid: (vi) intrinsic values, which did not focus on biological needs, but rather on abstract elements such as *truth* or *beauty* (see [101]). Nonetheless, this theory has been severely questioned: there are empirical evidence that show how human needs are not always progressively pursued in that hierarchical view (see [40]). For instance, social belonging has been suggested to be at the most basic level of human needs and not pursued only once physiological and safety needs are met [26].

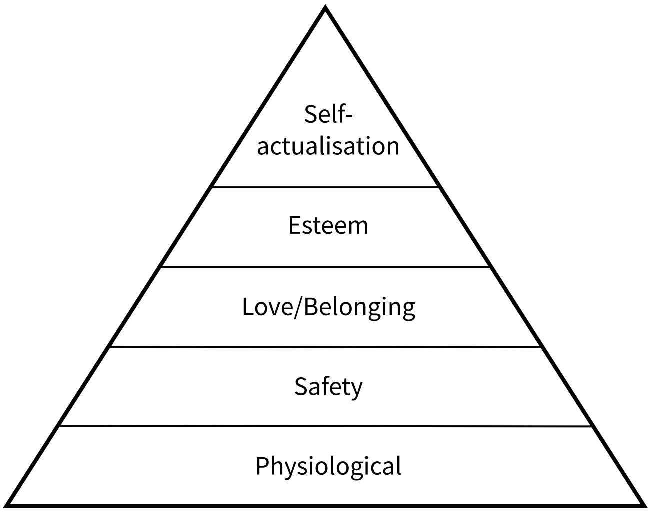


Figure 2.1: Maslow's hierarchy of needs (Source: [159])

Another theory is the one developed by Reiss. Reiss [203] proposed a theory of basic motives or desires, which were defined as reasons individuals hold for initiating and performing voluntary behaviour, thus possibly revealing their values and the meaning of their behaviour. He identified 16 motives or desires: (1) power; (2) curiosity; (3) independence; (4) status; (5) social contact; (6) vengeance; (7) honour; (8) idealism; (9) physical exercise; (10) romance; (11) family; (12) order; (13) eating; (14) acceptance; (15) tranquillity; and (16) saving.

Noteworthy progress in the research about the connection between values, needs and motivations has been done by Rokeach [207], Schwartz [226], and their colleagues. Rokeach [207] developed the Rokeach Value Survey to infer the value priorities of individuals, and distinguished between *instrumental values* (i.e. related to modes of behaviour) and *terminal values* (i.e. desirable end-states of existence, such as freedom and familiar security); and also between *individual values* (i.e. related to satisfying individual needs and self-esteem, such as honesty or courage) and *social values* (i.e. related to social demands, since *supra-individual entities* (e.g. society, organisations, etc.) "socialise the individual for the common good to internalise shared conceptions of the desirable").

Working on Rockeach's formulations, Schwartz postulated the Theory of Basic Values [226], which has been supported by substantial empirical evidence [221], becoming the mainstream (empirically-supported) value theory (see [104]). One of the main observations was that human values could be captured by a finite set and were consistent across cultures. Accordingly, Schwartz and Bilsky [225] suggested that values are cognitive representations of human needs, and identified three basic human needs: (i) the needs of individuals as biological organisms; (ii) the requisites of coordinated social interaction; and (iii) the survival and welfare needs of groups.

The Schwartz Theory of Basic Values defines 10 groups of *motivational values* according to the type of goal or motivational concern they pursue. These groups are graphically organised in a circle, in which those groups that are commonly exhibited together are positioned close to each other (Fig. 2.2). This spatial representation enables to easily identify "conflicting" values (that is, those groups that are in diametrically-opposed positions). The 10 Basic Values are: (1) power; (2) achievement; (3) hedonism; (4) stimulation; (5) self-direction; (6) universalism; (7) benevolence; (8) conformity; (9) tradition; and (10) security (Table 2.1).

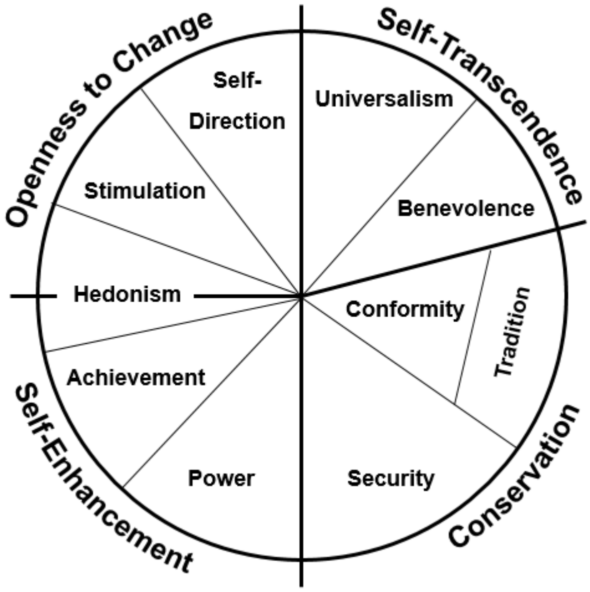


Figure 2.2: Model of the ten motivational types of value in the Schwartz Theory of Basic Values (Source: [219])

This theory has been used to study political behaviour (e.g. voting behaviour [227] and decision-making in public policies [37]), and even has been a subject of study in order to enhance their usability in public administration and policy studies (e.g. [270]).

Recently, Schwartz and colleagues [220] proposed to refine the Schwartz Theory of Basic Values in order to provide greater explanatory power and emphasising that values form a circular motivational continuum. In this spirit, they postulated 19 values sets (mostly, splitting the previous value sets into two different types). For example, they divided 'benevolence' into 'benevolence–dependability' (focusing on being a reliable and trustworthy member of the group) and 'benevolence-caring' (focusing on the devotion to the welfare of group members).

Table 2.2 provides an overview of all the theories reviewed in this section.

2.1.3 Values from a social perspective

Values can also be explored at a cultural level. As human beings are social beings, there are notions of common good from a social perspective (thus principles and standards that define the public interest of a community or society). In this spirit, the notion of values has been explored also for social goal-setting and decision-making in public affairs. Noticeably, this notion of value transcends individuals' motivations and relates to what is good and worthy from a supra-individual perspective (although it

Table 2.1: Value sets and their theoretical goal of the Schwartz Theory of Basic Values (Source: [219])

SELF-DIRECTION: Independent thought and action —choosing, creating, exploring (creativity, freedom, choosing own goals, curious, independent).	**STIMULATION:** Excitement, novelty, and challenge in life (a varied life, an exciting life, daring).
HEDONISM: Pleasure or sensuous gratification for oneself (pleasure, enjoying life, self-indulgent).	**ACHIEVEMENT:** Personal success through demonstrating competence according to social and cultural standards, thereby obtaining social approval (ambitious, successful, capable, influential).
POWER: Attainment of social status and prestige, and control or dominance over people and resources (authority, wealth, social power, preserving my public image, social recognition).	**SECURITY:** Safety, harmony, and stability of society, of relationships, and of self (social order, family security, national security, clean, reciprocation of favours).
CONFORMITY: Restraint of actions, inclinations, and impulses likely to upset or harm others and violate social expectations or norms (obedient, self-discipline, politeness, honouring parents and elders).	**TRADITION:** Respect, commitment, and acceptance of the customs and ideas that one's culture or religion impose on the individual (respect for tradition, humble, devout, accepting my portion in life, moderate).
BENEVOLENCE: Preservation and enhancement of the welfare of people with whom one is in frequent personal contact (helpful, loyal, forgiving, honest, responsible, true friendship, mature love).	**UNIVERSALISM:** Understanding, appreciation, tolerance, and protection for the welfare of all people and for nature (broadminded, social justice, equality, world at peace, world of beauty, unity with nature, wisdom, protecting the environment).

is true that contribution to the common good is also a basic individual motivation —see, for instance, *universalism* as a value group in the Schwartz Theory of Basic Values mentioned in the previous section). Therefore, it is necessary to explore another concept of values that is relevant for assessing the effects in a social system.

For individuals, values not only have to do with the individual conduct, but they also relate to beliefs about how the society should be. Individuals evaluate and reason about public affairs and social outcomes using values [111, 37, 135]. Hence, values have been notably covered also in political philosophy, policy sciences, and public administration, to discuss on value-driven decisions in public affairs. In other words, there has been extensive research aiming to define what constitute desirable society-level outcomes.

Many scholars have focused on eliciting universal sets of *social values* (also referred to as *societal values* or *public values*), constituting consensual expressions of what is desirable in society (see [39]). Accordingly, Bozeman [39] claims that such values provide "normative consensus about (a) the rights, benefits, and prerogatives to which citizens should (and should not) be entitled; (b) the obligations of citizens to society, the state, and one another; and (c) the principles on which governments and policies should be based".

With this in mind, policy-making has been described as "the process by which governments translate their political vision into programmes and actions to deliver desired changes in the real world" [256] —where "political visions" define the concept of common good. Policy-making is (in rather simple terms) a process to design and enact public policies, which are plans or programmes for coordinating the social activity towards a desirable social-outcome (for instance, establishing a tax on sugary drinks to reduce the consumption and thus contribute to the health of the community —that is, a collective

Table 2.2: Comparison of value theories focused on the individual

Value theory	**Focus**	**Categories**
Maslow [159]	Hierarchy of needs	(5): Physiological; Safety; Love and belonging; Esteem; Self-actualisation
Reiss [203]	Motives of actions	(16): Power; Curiosity; Independence; Status; Social contact; Vengeance; Honour; Idealism; Physical exercise; Romance; Family; Order; Eating; Acceptance; Tranquillity; Saving
Rokeach [207]	Desirable lifetime end-states and modes of behaviour	(18) Terminal values (e.g. Mature love; Equality; Freedom; Peace); (18) Instrumental values (e.g. Ambition; Love; Honesty; Obedience)
Schwartz [221]	Motives of action	(10): Power; Achievement; Hedonism; Stimulation; Self-direction; Universalism; Benevolence; Conformity; Tradition; Security
Schwartz [220]	Motivational continuum of action	(19): Power-resources; Power-dominance; Achievement; Hedonism; Stimulation; Self-direction-action; Self-direction-thought; Universalism-tolerance; Universalism-nature; Universalism-concern; Benevolence-caring; Benevolence-dependability; Humility; Conformity-interpersonal; Conformity-rules; Tradition; Security-societal; Security-personal; Face

response to a collective problem). This means that policy-making and public policies involves value-laden decisions, that not only reflect the standards that ground the notion of common good and public interest, but also the hierarchy of those conflicting values [37]. [2].

Given that these values concern the society at large and not only the individual holding it (see [135]), these values may be opposed to the private self-interest of individuals [39] (e.g. elder people accepting to pay higher taxes for better basic education, although they are not directly benefited from improving this public service). In spite of assuming that individuals may hold different social values in comparison to each other, they are often shared among social groups (see [135]). As mentioned previously, there is a reciprocal interaction between the individual and the group [110]: the group influences the individual through its shared values just as the individuals influences the group through its subscription (or not) to these values. This influence may also appear in companies and organisations, and not only in societies [236]. In this light, Kenter *et al.* [135] distinguished between *shared values* and *social values*: the former are values that were shared among individuals, and the later are those values that concern the society at large.

---

[2] Indeed, values have to do with this view of (Aristotelian) *politics*, understood as the art of organising public affairs whose purpose is to achieve a good social outcome (also, common good or public interest). Notice that another usual understanding of *politics* is about contesting competing views of good social outcome and appropriate organisations of public affairs, emphasising the interaction between competing stakeholders.

Cognitive view of values in policy decisions

In contrast, other scholars explore these values from a cognitive point of view (and not from a normative approach for guiding policy-making). For instance, they explore how values are involved in decision-making of public servants [37, 270] and managers [133] in the context of their work —that is, the standards that eventually determine their decisions. In this sense, *public values* are beliefs that decision-makers invoke to justify their choices in public affairs on the behalf of the community [271].

Thus, *public values* act as guiding principles in political decisions. Thacher and Rein [253] described values in public policies as "the ultimate ends —the goals and obligations that policy aims to promote as desirable in their own right, not just as means to some other objective" [253]. However, the understanding of what constitutes *public value* (i.e. the worthiness of the social outcome or state of affairs) arises from evaluations made by individuals according to their value hierarchies, which leads to the preference of particular social outcomes (see [168]). Interestingly, public decisions usually pose ethical dilemmas among desirable outcomes and conflicting values [136, 270, 37] (e.g. privacy vs. security). Besides, values are involved in making sense of a collective problem (which was previously presented as *perception*). As Stewart [247] claimed, "policy design has a value-based component because the ways we attempt to change or influence behaviour depend on, in turn, beliefs about the reasons for that behaviour".

Based on this perspective, any public policy is a "value-driven plan of action", which describes a collective problem that requires intervention, defines a *better* state that should be attained, and establishes some governance mechanisms to do so. That is, public policies have to do with values (and priorities) and the economy of public resources [37]. In the most simple terms, public policies declare *ends* (i.e. what states of affairs are desirable) and *means* (i.e. how this state is going to be achieved, by intervening a social system).

Cognitive view of values in political behaviour

Values have been addressed to explore the political cognition of individuals, which is growing attention nowadays.[3] Core beliefs and human values are basic criteria for political evaluations within policy domains (i.e. policy preferences, performance judgements, candidate assessments, etc.) [83, 94], eventually influencing voting behaviour [227] and influencing the social acceptance of public policies [269]. In this view, it has been suggested that ideology is the way values are expressed and debated in political life [37].

Moreover, some scholars study how individuals use socio-cognitive heuristics to make sense of the political world [3]: for instance, by voting for the candidate who most aligns with their ideological worldview, by relying only on information from aligned sources, usually failing to link cause and effects in policy decisions, and being subject to framing effects [144]. In this light, Achen and Bartels [3] contrast the actual political behaviour of individuals with the ideal theories of democracy that consider citizens as rational, well-informed voters, concluding that social identity and political loyalties are even more decisive than rational policy preferences. In fact, research has been done on linking values and personal identity, which in turn leads to the constructions of group- and value-identities [108, 152].

Values, public services and governance

Paraphrasing Moore [168], the contribution to *public value* (i.e. the worthiness of a state of the world) justifies taxation in order to invest in public services and public policies, thus serving citizens of the

---

[3] See, for instance, the call for experts from the Joint Research Centre, aiming at collaborating to improve the understanding of values in policy-making and political decision-making: `https://ec.europa.eu/jrc/en/enlightenment/call-experts-science-of-values`

society. In this sense, *public values* represent a compelling story that connects the public services and public policies to its worthiness [177]. For example, water sanitation is an essential public service in cities because it contributes to public health by protecting urban population from waterborne diseases.

As values are involved in decisions in public affairs, they influence the way the society is governed and the way public services are delivered. In the same vein, citizens also have beliefs about how the society should be. Consequently, citizens want governance and public services to be provided according to their values, their political preferences, and their ethical expectations [168].

These demands can be expressed in a political arena (i.e. citizen), and not only in markets by making different choices (i.e. consumer), because one's preferences as a consumer may differ significantly from one's preferences as a citizen. As Moore [168] points out, "politics is the answer that a liberal democratic society has given to the (analytically unsolvable) question of what things should be produced for collective purposes with public resources". With this in mind, citizens will judge the *value* of those companies —especially, those that manage public services— against their expectations of justice, fairness, efficiency and effectiveness [168].

However, it is true that citizens do not always have the necessary information to make certain decisions. For this reason, it has been suggested that citizens should not make technical decisions on their own (see [70]). This means that although it is completely legitimate for them to make decisions and express their views on the values that should guide the management and provision of a public service, they should be fully informed and consider technical advice by experts (e.g. engineers, scholars, etc.).

Theoretical approaches and taxonomies of social values

Many scholars have proposed inventories of *public* or *social values*, with different foci. (i) Some sets are centred on providing theory to analyse culture and socio-political changes (i.e. progress lines of societies); (ii) others are addressed to (objectively) assessing management and public administration (i.e. normative or prescriptive political views); (iii) and, finally, others are proposed to study the beliefs of individuals about the organisation of the society (i.e. descriptive political views), focusing specially on public servants and their decisions in public affairs. Although the research questions of these approaches differ, it is clear that the content of the sets of value items is mostly coincident.

One example of the sets of the first category is Inglehart's theory [119], who, working on Maslow's theory, related values with political shifts in societies. Their work aimed to demonstrate that post-industrial societies, once ensured their economic security, focused on more transcendental goals like the protection of the environment. He defined two main sets of values: *materialist*, related to the lower levels of the Maslow's pyramid (i.e. stability, security, social order), and *postmaterialist*, related to the upper levels (i.e. citizen involvement, freedom, equality, etc.). In the same vein, *materialist* values were related to security and power in the Schwartz's theory, and *postmaterialist* values were related to universalism (which are opposed in the Schwartz's spatial representation (Fig. 2.2)).

Several examples can be found in the second category. Although there are many taxonomies, they concur with reflecting the diverse duties and responsibilities in governance [151].

Kernaghan [136] divided values into four categories: *ethical values*, such as honesty and integrity; *democratic values*, such as legality and accountability; *professional values*, such as efficiency and innovation; and *people values*, such as tolerance and compassion.

Hood [114] proposed three types of administrative values: *theta values*, that refer to honesty and rectitude; *lambda values*, that relate to resilience and robustness; and *sigma values*, that are about efficiency and frugality.

Jørgensen and Bozeman [128] reviewed public administration literature and identified a list of 72 values that they organised based on the seven aspects of public administration the value affects: (1) Public sectors' contribution to society (e.g. *public interest*, *human dignity*, *regime stability*, etc.); (2) transformation of interests to decisions (e.g. *democracy*, *citizen involvement*, etc.); (3) relationship between public administrators and politicians (e.g. *accountability*, *responsiveness*, etc.); (4) relationship between public administrators and their environment (e.g. *listening to public opinion*, etc.); (5) intra-organisational aspects of public administration (e.g. *reliability*, *timeliness*, *effectiveness*, etc.); (6) behaviour of public employees (e.g. *moral standards*, *professionalism*, *ethical consciousness*, etc.); and (7) relationship between public administrators and the citizens (e.g. *equal treatment*, *protection of the rights of the individual*, etc.).

Regarding values as decision-drivers, Witesman and Walters [270] studied the beliefs of public servants, who made decisions on behalf of many individuals and whose values were reflected in their choices. The authors compared universal *public values* (e.g. Jørgensen and Bozeman's inventory [128]) and *motivational values* (e.g. Schwartz's basic values [221]), and concluded that (i) the current public values literature lacked the micro-theoretical grounding to provide an understanding of the psychological dimensions that impact the individual decisions of individual public servants; and that (ii) the Schwartz's theory —the most consolidated theoretical proposal on values— lacked social values that accurately capture the preferences of societies in aggregate and that are very likely to be invoked by individual decision-makers in making decisions in their roles as public servants.

Consequently, Witesman and Walters [270] postulated the existence of *public service values* as "the subset of social, professional, ethical, and other values that are related directly to a person's role as a public servant, and would be acknowledged by that public servant as reasonable, legitimate, and relevant in carrying out the functions of a given position in the public sector". They proposed a modified version of the Schwartz Theory of Basic Values, to shift from values involved in the preferences of individuals to values in the preferences of societies at large (Fig. 2.3). This work considered 15 theoretical value groups —for which they split Schwartz's *universalism* into six different sets (Table 2.3): (1) Community/Responsiveness; (2) Transparency/Citizen influence; (3) Social Justice; (4) Neutrality; (5) Stability; (6) Equity; (7) Benevolence; (8) Tradition; (9) Conformity; (10) Security; (11) Power; (12) Hedonism; (13) Achievement; (14) Group interest; and (15) Stimulation. This theory has been applied to study actual decisions by public servants with some changes [271, 272]. Alas, as the authors warn, it is still a novel approach and it should be further explored with more empirical evidence. Besides, its "spatial representation" has been severely questioned (as it does not provide the same meaning as does in the Schwartz's theory) [272].

Also Schwartz [224] worked on further exploring the value items of the *universalism* value set (e.g. social justice, equality, tolerance, etc.) to observe their constitution as moral social values that are involved in many policy issues (e.g. immigration).

Table 2.4 provides an overview of the value models reviewed in this section.

2.2 Policy assessment

Values are involved in designing policies, either evaluating end states and choosing means of action. Since these choices have significant economic, social or environmental impacts, it is advisory to conduct a policy assessment. A policy assessment is an explicit statement of the expected environmental, social and economic impacts, describing who are affected by the intervention and how, and providing some consultancy [76].

In plain terms, a policy is a plan of action in order to reach a state deemed desirable or achieve some goal. Public policies aim at addressing what has been defined as a collective problem [65] and pro-

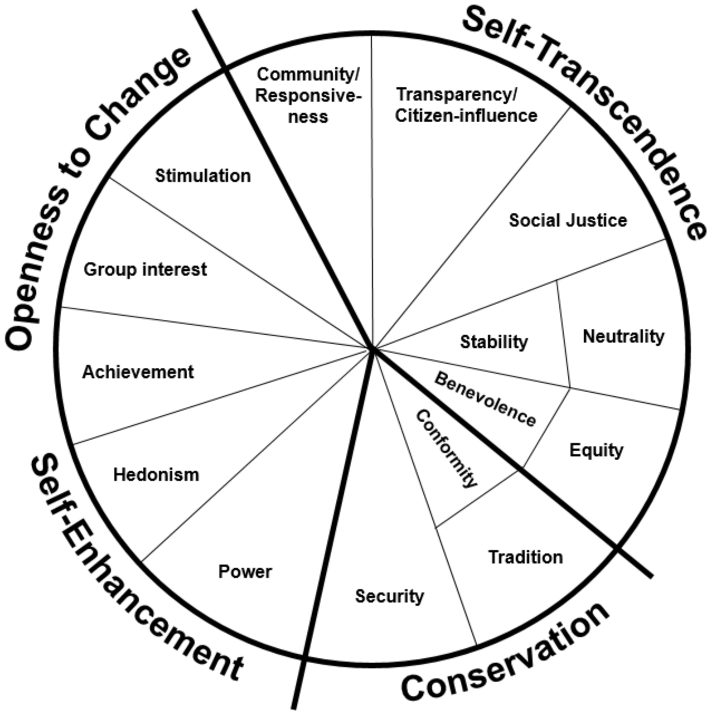


Figure 2.3: Model of the Witesman and Walters' Public Service Values (Source: [270])

ducing a desirable society-level outcome [168, 129]. According to May [160], public policies "typically contain a set of intentions or goals, a mix of instruments or means for accomplishing the intentions, a designation of governmental or non-governmental entities charged with carrying out the intentions, and an allocation of resources for the requisite tasks".

2.2.1 Policy-making cycle

Public policies also have to do with power. According to Botterill and Fenna [37],"in the real world of policy decision-making, policy choices emerge from the interplay of a variety of actors and advocacy coalitions who represent and seek to promote and defend different sets of relevant values".

The policy-making cycle is often described as a linear process that includes agenda-setting, design, implementation, application (including enforcement), evaluation and revision (i.e. maintain, redesign, or terminate the policy). However, policy-making and its stages are far more complex and uncertain than a linear process [47, 214] and multiple theories have been postulated in this regard (see [65, 37]). Some of the most salient features that explain that complexity are that (i) while design is usually made without enough information, decisions may have substantial impact, which is not always direct and it is often difficult to quantify or even foresee; (ii) several factors and actors are at play, their interdependence may not be evident, and they may be out of control the of any actor, but they still determine the policy outcome; (iii) multiple social needs are at stake, but their meaning may depend on the perspective collective problems are approached with; (iv) multiple stakeholders are involved, who have competing values and interests, advocate for them in political arenas, ad mobilise diverse resources (e.g. economic, information, etc.) [65] —for instance, occupational groups, religious groups, etc.; (v) stakeholders use many forms of knowledge besides scientific evidence when making decisions like habits and conventional wisdom [47, 38, 271].

Table 2.3: Value sets and their operational goal of Witesman and Walters' Public Service Values (Source: [270])

SOCIAL JUSTICE: Seeking justice and promoting the welfare of society as a whole (public interest, human dignity, sustainability, common good, social cohesion).	**NEUTRALITY:** Achieving freedom from bias; allowing actions and information to pass through people without impacting them (government neutrality, political neutrality, impartiality).	**EQUITY:** Distributing social benefits without discrimination or favouritism; support of systems and actions that promote fairness and equality for individuals and groups (equal treatment, social justice, equity, fairness).
BENEVOLENCE: Promoting the welfare of people with whom one has personal contact (responding to need, advocacy, friendliness, loyalty).	**TRADITION:** Upholding customs derived from social institutions (rule of law, regime dignity, loyalty).	**CONFORMITY:** Refraining from potentially unacceptable action (honesty, moral standards, ethical consciousness, integrity, legality).
SECURITY: Avoiding danger or deterioration of quality (regime stability, national security, security of information).	**POWER:** Achieving control or dominance over people and resources (being in charge, decisiveness, prestige, authority).	**HEDONISM:** Achieving personal pleasure and gratification (good working, self-interest).
ACHIEVEMENT: Achieving competence or success according to social standards (reliability, productivity, effectiveness, efficiency, professionalism, competence, responsiveness).	**GROUP INTEREST:** Promoting the welfare of a group or segment of all people (protection of minorities, voice of the future, stakeholder value).	**STIMULATION:** Pursuing excitement and challenge (innovation, challenge, adaptability, risk readiness).
COMMUNITY/RESPONSIVENESS: Pursuing mutual benefit of two or more individuals or groups; sharing of power and prestige to the mutual benefit of two or more individuals or groups (cooperativeness, compromise, collaboration).	**TRANSPARENCY:** Allowing access to information; making information and actions available for scrutiny (transparency, openness, objectivity).	**CITIZEN INFLUENCE:** Upholding influence of individuals and groups over the institutions and policies that affect them (citizen involvement, local governance, dialogue, democracy, will of the people, listening to public opinion).

2.2.2 Policy design and assessment

In this dissertation the focus is on the *ex ante* policy assessment, without discussing in depth policy-making theories and how stakeholders interact in the cycle. Thus, although policy processes involve the interaction of multiple social actors that shape the actual policy, this dissertation is focused on standalone users who use computational models for social simulation in order to enrich their knowledge and introspect about their policy decisions.

The interest is on how to model the domains to be intervened (including stakeholders) and the policies to address a collective problem, in order to perform an assessment of those policies. In fact, the European Commission refers to this process as Impact Assessment (IA), and considers it necessary when the expected economic, environmental or social impacts of actions are likely to be *significant* (see [76]). With this in mind, social simulation is considered a suitable methodology to perform such *ex ante* assessment in a more sophisticated way.

In a general view, according to the European Commission an Impact Assessment should include the following stages:

Table 2.4: Comparison of value theories focused on the society

Value theory	**Focus**	**Categories**
Inglehart [119]	Hierarchy of societal goals	(2): Materialist values (e.g. security, tradition, stability); Postmaterialist values (e.g. freedom, self-direction)
Kernaghan [136]	Public administration and public management	(4): Ethical values (e.g. Integrity; Accountability; Honesty); Democratic values (e.g. Rule of law; Neutrality; Legality); Professional values (e.g. Efficiency; Excellence; Innovation); People values (e.g. Caring, Benevolence, Humanity)
Hood [114]	Public administration and public management	(3): Sigma-type values (e.g. Frugality; Competence; Efficiency); Theta-type values (e.g. Honesty; Fairness; Mutuality); Lambda-type values (e.g. Reliability; Robustness; Adaptativity)
Jørgensen and Bozeman [128]	Public administration and public management	(7): Contribution to society (e.g. public interest); Transformation of interests to decisions (e.g. democracy); Relationship with politicians (e.g. accountability); Relationship with environment (e.g. public opinion); Intra-organisational aspects (e.g. timeliness); Behaviour of public employees (e.g. moral standards); Relationship with citizens (e.g. equal treatment).
Witesman and Walters [270]	Motives of decisions in public affairs	(15): Community & Responsiveness; Transparency & Citizen influence; Social Justice; Neutrality; Stability; Equity; Benevolence; Tradition; Conformity; Security; Power; Hedonism; Achievement; Group interest; Stimulation

Definition of the problem and boundaries of the system

The first step consists in abstracting the reality and drawing the boundaries around the domain. It should include, at least, an analysis of the situation and context, stressing the purpose and scope (i.e. spatial and time scales), and a description of the stakeholders involved and their motivations. In other words, a messy problematic situation is converted into a structured well-defined problem, affording to design policies and hence address collective problems [115]. Given this definition of the problem, one can reason about the *ends* and the *means*.

In general terms, significant ethical and political assumptions —and not only scientific— are made in the processes of abstracting the reality to build these *policy models* (see also Chapter 3). These involve value judgements that ground the definition of the problem, the choice of alternative instruments suitable to address the issue, and the evaluation system that will be used to assess the social outcome. Accordingly, defining the problem correctly is the most important stage, as it will determine the policy design and thus the policy effectiveness.

Accordingly, the policy model presents a specific definition of a collective problem and links it with those who are held accountable (either as cause or solution of the issue). For instance, an excessive water use approached as collective problem may be represented in different ways that involve distinct moral considerations. Representing such issue as a matter about the diffusion of efficient technologies in a market and their adoption by households has a significant underlying meaning: in this case, the policy model declares households as target group —portrayed as consumers— and frames the solution as a market-choice. What is not represented in the model may be even more important; in this example, industries are not the target group, and the solutions are not, for instance, taxation of the profit they make from using water as input of their activities.

This is why it is said that policy models reflect paradigms, which are taken-for-granted descriptions and theoretical analyses that constrain the range of alternative policy options [50,125]. Pahl-Wostl [181] states that “paradigms strongly influence meaning, understanding and perception of reality and of problem situations, how boundaries are delineated, and how the space for identifying problems and developing solutions is determined”. Alternatively, Campbell [49] defines paradigms as “cognitive background assumptions that constrain action by limiting the range of alternatives that policy-making elites are likely to perceive as useful and worth considering”. In essence, paradigms establish worldviews and theoretical frameworks that constrain the range of policies to be considered when addressing public issues, as economy or social welfare [50, 125]. It has been suggested that these paradigms are supported by language and discourse, contributing to form “mental structures that shape the way we see the world” [142].

Policy objectives

The second step focuses on establishing and describing clearly what objectives should the intervention is meant to achieve, thus defining somewhat the desirable states to be attained. Furthermore, these objectives constitute the benchmark that will be used for *ex post* evaluations of the policy —or *ex ante*, if simulation is used.

In order to assess the policy, policy-makers and stakeholders may rely on some performance indicators that stand for those end states and make objective the assessment of improvements [112]. Indicators may be quantitative or qualitative, and more or less direct about the value they pursue.

These indicators are specific for a particular context and domain, and should involve aspects of accuracy, time and cost of collection and processing. Furthermore, it is convenient to remark that the selection of indicators to capture values is not a *neutral* decision, and that should not become values on their own (but rather, they are projections of those values onto an element that can be measured).

Policy options

The third step identifies which alternative options are available to reach the policy objectives. Besides being consistent with the definition of the problem, such options should take into account that target groups and field staff must accept and support them at some degree [160]. Usually, public policies aim to produce a behavioural change on target groups so as to drive the system towards a desirable state.

Alternative approaches and combinations may be available to address the problem: financial instruments (e.g. grants, subsidies, loans, insurance, etc.); economic instruments (e.g. prices, taxes, tributes, etc.); regulatory instruments (e.g. laws, standards, etc.); informational/persuasive instruments (e.g. messages, campaigns, etc.); social instruments (e.g. networking, distribution channels, etc.); and resource/service supply instruments (e.g. infrastructure, etc.), among others.

Also, *ex ante* assessments should explore the feasibility and certainty that an action or plan will produce that desired outcome in the time and spatial scales of interest.

Policy effects and evaluation

The fourth step is intended to evaluate the expected results of the policy in the social system. Assuming that simulation is used to carry out an *ex ante* assessment, this process is done by using a model of the system and its computational simulation. Nonetheless, a plan for monitoring the evolution in the real social system should be elaborated, as it will be necessary for continuous evaluation and revision.

Values are involved in the evaluation of the resulting states of the system, as they establish the standards and criteria that express the worthiness of some state. With this in mind, diverse methods are used to evaluate the outcome, such as cost-benefit analysis (CBA) or multi-criteria decision analysis (MCDA) (see [171, 195]). These methodologies support the comparison between options by assuming that values can be projected onto some indicators (a *consequentialist* view) and can be compared and aggregated, allowing value trade-offs (a *commensurable* view) (see [116]).

Interestingly, a baseline scenario (i.e not intervening at all) should be always considered as a benchmark for evaluation.

2.3 Artificial intelligence and agent-based simulation

Given the profound social consequences (e.g. ecological, social, economic, etc.) that public policies may have, the complexity of social systems, and the limited cognitive capacity to consider all possible actions and effects in social systems, it is convenient to assess policies prior to their enactment. Artificial Intelligence (AI) may provide powerful tools and methodologies for this matter.

Simulation is the imitation of a real-world process or system over time and can contribute to policy assessment without disturbing the real social system and committing resources [23]. Simulation tools can be used on social system models to explore the effects of alternative policies, as well as identify counter-intuitive situations that should be avoided. With this aim, policy simulation requires models that capture the social dynamics of the domain of interest, in order to appropriately inform policy-making. That is, it is necessary to build appropriate models of these collective problems, in order to apply computational tools that help experts to address these better in the real world.

2.3.1 Agents and socio-cognitive technical systems

Technically speaking, *agents* are entities that are able to perceive the environment in which they are situated and act autonomously according to an internal model of rationality [212]. In this light, a focus of AI consists of exploring and developing faculties for artificial agents (i.e. computer agents) in order for them to perform a certain function or task —thereby creating *artificial intelligences*.

Besides, although the field of AI has focused traditionally on individual intelligence, it has been explored social phenomena (see [22]), especially considering the field of multiagent systems. Indeed, human intelligence comprises socio-cognitive aspects because of the social nature of human beings. Thus, AI can be used to study the coordination of systems constituted by social agents, in order to engineer technologies and methodologies that support the reasoning and decision-making on collective problems.

On this basis, the notion of socio-cognitive technical systems (SCTS) arises. SCTS are social coordination systems [12] that articulate the interactions of autonomous agents that are socially rational [174]. They distinguish two first-class entities: a *social space* where all interactions take place, and *agents* that interact autonomously within that environment. This view is particularly useful for policy-making, because it sets the distinction between the agent, whose behaviour is motivated by values, and the social space, that must be governed to preserve social values.

2.3.2 Agent-based models and social simulation

Agent-based models (ABM) are a type of computational model whose purpose is the simulation of actions and social interactions of individuals or groups within an artificial environment —which is also known as social simulation. The field of agent-based social simulation (ABSS) is used to develop greater understanding of the complexity of interconnected societies, providing a useful approach for diverse scientific disciplines such as biology, sociology or economy. Not surprisingly, the exploration of sociological aspects by means of computer models is known as *computational social sciences* [59].

This approach is particularly useful for studying the social behaviour of multiagent systems, being able to simulate interactions of multiple heterogeneous agents at the micro-level that eventually produce social phenomena at the macro-level. In this sense, ABMs require agents to be provided with an artificial intelligence that makes them exhibit a realistic social behaviour in the domain of interest. Hence, the models constitute a society of artificial agents that emulates the behaviour of the real social system [90]. Following this vision, and depending on the purpose of the model, it will be necessary to include multiple submodels:

- Biophysical models (e.g. ecological dynamics, population dynamics, agricultural production, etc.) (e.g. [16]);
- Models of human behaviour (e.g. reasoning, ethics, economic behaviour, cognitive biases, etc.) (e.g. [235, 118]);
- Models of social topology and interaction (e.g. influence, peer pressure, social networks, communication, etc.) (e.g. [103]);
- Economic models (e.g. production of goods, consumption of resources, purchase-sale, diffusion of innovations, etc.) (e.g. [31]);
- Spatial representation models (e.g. grids, topographic maps, etc.) (e.g. [189]);
- Political models (e.g. argumentation, discourses and influence, framing, etc.) (e.g. [102]).

Computational models for policy-making

Agent-based social simulation has been acknowledged as a useful methodology to support policy-making [91, 124, 131]. In this sense, models can be used to infer how policy interventions would perform in the actual societal system from their effects in a simulated society.

However, one of the main barriers of this approach is to build realistic but simple societal models. Many social and psychological theories are not expressed in a way that their implementation in computer models is simple, and neither they are referred to the behaviour of large masses of individuals in complex systems. Hence, many scholars are concerned about social-simulation models being artifices and erroneous, since their hypotheses and assumptions have been built to fit the data (overfitting) and have not been properly validated [87]. So, what utility can be derived from social simulation?

On the one hand, although models do not reflect *exactly* actual socio-cognitive processes, this does not mean they are not *realistic*. Therefore, even if the model is “artificial”, the simulation of realistic micro-interactions in a complex system can be useful, as they may result in interesting situations that may expose the need for more in-depth investigation. This type of results generate scientific questions: “*This is unexpected. Why does it happened? Does it make sense?*” (see [145]).

Accordingly, these tools do not really predict, but rather they clarify the dynamics of the system under diverse conditions, which can support policy assessment [6]. Actual prediction in policy domains is practically impossible, as they are complex due to the interdependence between environmental, economic, social, political and cultural aspects [91, 53]. This view assumes that the long-term macro-behaviour is consequence of many short-term micro-decisions, being the former likely to be predicted in rough terms. Hence, prediction in agent-based models refers to “the ability to reliably anticipate data that is not currently known to a useful degree of accuracy via computations using the model” [72].

On the other hand, modelling is useful in that it is a research process. Every decision made when building the model contributes to understanding the phenomena and may lead to consider the need for more detail. Furthermore, it structures the problem in a way that it can be easily updated with more appropriate sub-models. In sum, the value lies not so much in the model itself, but rather in the conclusions drawn from simulations and in the exploration of a real-world analogy.

One should understand that all models are simplifications, and therefore no model will be a complete representation of reality. In fact, this is what helps to better understand phenomena; hence, modelling always faces a trade-off between simplification and sophistication [248, 87].[4] Even under social complexity, individuals continue to use models —mainly, mental models— to make decisions. Imagining hypothetical situations to develop anticipation is quite common. With this in mind, oversimplified models and heuristics may be used to make decisions about public policies, which have consequences on many people. Many models admit only a certain level of complexity, besides being biased and opaque, so that they can generate erroneous predictions and eventually they can misguide decisions —and no one can explore and question that misguiding knowledge.

Interestingly, agent-based social simulation is a quite novel scientific field, and its actual impact on policy-making is still limited. Although promising, it also has grown several concerns, mainly because of the risk of backfiring if used without proper precaution [19, 87, 71] and the ethical considerations when modelling the policy problems (see Sec. 2.2.2). Thus, decision-makers and modellers should be aware of the assumptions and limitations since these tools may be given substantial influence in policy processes (see [138]). One of the major challenges in modelling agent-based models is data scarcity and having a proper representation of bounded rationality in the domain and context under study, for which it is recommended to work with domain experts [117].

2.3.3 Values in Artificial Intelligence

Collective problems in multi-agent systems have been approached as a matter of changing individuals’ behaviour. The focus has been on designing a governance system that regulates the social space (considering institutions, norms, etc.), while taking into account the individuals within (their interests, traits, acceptance, etc.) [75, 255]. Hence, much of the work in multi-agent systems focus on the interaction of the governance system and the agents involved. Nevertheless, values have not been explicitly considered in such approach, in general terms.

One strategy to include values has been to sophisticate utility functions, where the value is represented by a mathematical expression and weighted differently depending on the sensitivity of the agent [163]: for instance, the utility with regard to *fairness* may increase when some benefit is split equally among the agents of the environment. This approach is easily implemented for those values that have an immediate numerical translation (e.g. efficiency, fair-distribution, wealth-accumulation, etc.), but it may be far more difficult for values that are more abstract but important from a social point of view (e.g. social justice, security, social recognition, honesty, etc.). Besides, although this approach is practical, it may be seen as oversimplifying for many stakeholders [1]. It is assumed that rationally-bounded agents are more likely to use unspecified non-formalised evaluation frameworks, solving value conflicts by invoking some form of satisficing combinations [235] (and not some formal utility function). Moreover, for the sake of simplicity, these works usually consider that all agents in the society understand the values the same way, only differing the weights they give to the different values.

---

[4] These issues are condensed in the aphorism “all models are wrong, but some are useful”[5].

Thus, little research has been done to consider values with a socio-cognitive approach in agent-based simulation, despite its potential to study social behaviour (see [110]) and structure social coordination decision-making (see [134]). Some recent works included values as drivers of behaviours so as to consider economic, environmental and social factors [246, 106], emphasising the fact that working with abstract values in computational models requires unavoidably to translate them into operational terms (e.g. particular actions and indicators), which inevitably makes them subject to strong discussion.

Apart from these constructs in social modelling, values have been covered in argumentation [267, 20, 28] and normative systems [147] as well. Some scholars have also explored values in ICT-based policy-making, but focusing on stakeholders-model interaction processes (e.g. trust in the model, ownership of the model, etc.) rather than an element in the computational model (see [24, 167]), which is more related to how values steer the design and creation of technological artifacts (i.e. *value-sensitive design* [195, 196]).

Part II

Conceptual framework

Chapter 3

Foundations

Although policy-making is a complex cycle that involves diverse stages (such as negotiation and enactment), this dissertation is focused on policy assessment. Policies intervene in a social system in order to achieve a *better* state of affairs (that is, producing a *desirable* effect in the system). The key responsibility of policy-makers is to design the intervention in the social system: on the one hand, the policy-maker defines what constitute those deemed *better* state of affairs, hence declaring the policy ends; on the other hand, choosing the means that are likely to achieve the improvement of the state of affairs (which are usually a combination of actions that stakeholders are encouraged either to take or to avoid, implemented with instruments like norms, incentives, or persuasive messages). Moreover, policy-makers must consider that stakeholders in the social system act in accordance to their own interest and hence they may make the intended change more or less successful.

To address complex collective problems, the social system is abstracted into a model, translating a deemed problematic situation into a structured well-defined problem that allows to seek alternative solutions methodically [115]. The construction of these models include both factual knowledge of diverse areas (in particular, cause-effect relations in the system), but also ethical knowledge, which means that notions of good, rightness, and relevance (i.e. values) are also reflected. This underlies one main premise: that policy-making is a value-driven process.

These abstractions of the social system can be used to generate computational models for simulation, that allow to assess the effects of different policies *ex ante*. This reveals another main premise: that agent-based simulation is an appropriate methodology to support policy-making (especially, policy assessment and policy design). However, values must be conveniently considered when building the policy models that will support these policy-making stages.

For this reason, it is convenient to develop a conceptual framework that presents the policy-problems in a clear way, taking the role of values into account, and facilitating their computational approach. To this end, the foundations and assumptions that narrow the problem and support the conceptual framework (see Chapter 4) are exposed in detail here.[1]

3.1 Value-driven policy-making

Policy decisions have to deal with complex situations for which there is neither a simple nor a single correct solution (i.e. "wicked problems"). As a consequence, policy solutions can only be judged by the standards of better or worse (i.e. values), thus revealing a political nature [162].

---

[1] This chapter does not propose another scheme for policy assessment (which has been mentioned mentioned, broadly, in Chapter 2). Instead, much emphasis is placed on identifying the stages of the "(value-driven) cognitive process" through which the *problem* that is to be addressed with a computer system is *designed*.

For these reasons, policy decisions, like many other decisions, involve both factual and ethical questions, as Simon [237] pointed out (Fig. 3.1) (see also [185]):

(a) a *scientific question*, to know what will happen once an action is performed, (concerning beliefs of objective facts about the system and its dynamics); and
(b) an *ethical question*, to elucidate whether the outcome is desirable or not, or merely to determine what aspects deserve attention (concerning ethical standards or moral principles, to *improve* the state of affairs).

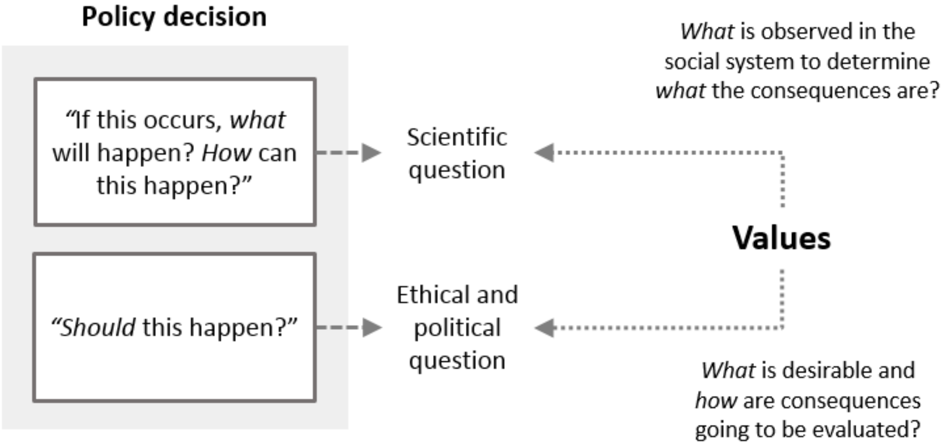


Figure 3.1: The two type of questions involved in a policy decision

Values are involved in creating the policy model that will support decision-making, which includes notions of how the state of affairs ought to be. Above all, values play a role in determining what is *relevant* of the system, and therefore deserves attention and monitoring with respect to a particular collective problem. Hence, both ethical and factual knowledge are present when defining the policy model (including relevant stakeholders, appropriate interventions, and the characteristics of the domain).

With this in mind, **values are present at multiple levels concerning policy decisions** (Assumption 1). First, values are involved in defining the normative problem (that is, that current state of affairs that *ought to be improved*) and its possible solutions (i.e. those *better* state of affairs to be achieved). Second, values are present when assessing the desirability of a state of affairs, but also affect the judgements about how well an action or plan will produce that desired outcome in a given time horizon (i.e. feasibility), since they guide the abstraction of the relevant part of the world and support beliefs that "complete" the lack of scientific evidence (see [2]). Third, it is clear that the multiple stakeholders involved in a collective problem may hold different values and interests, meaning that they may have particular understandings about how to deal with these problematic situations and how to judge the outcomes, thereby guiding their behaviour and their political appeals.

3.1.1 Policy simulator systems as a socio-cognitive technical system

Socio-cognitive technical systems (SCTS) are social coordination systems [12] that articulate on-line interactions of autonomous agents that are socially rational [174]. They are composed of two first-class entities: a *social space* where all interactions take place, and the *agents* that interact within that environment. The *agents* are autonomous entities who have a socio-cognitive model (that is, they base their actions on an internal decision-making model that takes social aspects into account). The *social space* is an environment with fixed ontology, restricted to the domain of interest, that at any given time it is in a *state* (i.e. all its variables are instantiated), which can change only as a result of an action or event that complies with the regulations of the system ant that is shared by all *agents*.

These theoretical ideas are useful to approach policy models for simulator systems. In particular, **agent-based policy simulator systems can be implemented as socio-cognitive technical systems** (Assumption 2) because:

(i) the *improvement* of the *state* of the *social space* is the end of policy-making;
(ii) the *social space* can be regulated in order to achieve that *improvement* of its *state*;
(iii) autonomous (and heterogeneous) *agents* act within the *social space* and change its *state*, who are subject to regulations and other instruments (which may affect their behaviour);
(iv) *agents* have their own goals and their socio-cognitive models (that include motivations, values, beliefs, etc.);
(v) there may be *events* that can also change the *state* of the *social space*;
(vi) the system is situated in a particular *domain* (e.g. water, energy, agriculture, etc.), and has a developer who designs it and an owner who uses it.

That said, it is useful to decompose SCTS in three views (known as the "$\mathcal{WIT}$ *trinity*"), according to Noriega *et al.* [174]:

(a) the fragment of the world that is relevant for the system "$\mathcal{W}$";
(b) the institutional representation of the conventions and regulations that define the system "$\mathcal{I}$"; and
(c) the implementation of this institutional world that creates the on-line version of the relevant world "$\mathcal{T}$".

In the case of policy simulators, these three views correspond to (a) the simulated world, (b) the formal model of the world and (c) the computational implementation of this formal model (see Sec. 3.1.2). Accordingly, these views are interrelated in such a way that an attempted action in the simulation modifies the *state* of the world "$\mathcal{W}$" if and only if that action complies with the conventions established in the formal model "$\mathcal{I}$" and can be processed in "$\mathcal{T}$" (i.e. the conventions are properly coded in the software model).

In practice, the institutional specification "$\mathcal{I}$" is achieved by instantiating a "*meta-model*", that includes *ad-hoc* constructs and data structures to represent key distinctive features for a certain purpose. In this dissertation, the conceptual framework supports the building of policy simulator systems, which are therefore approached as a sub-class of SCTS —henceforth coined as ***v**alue-**d**riven **p**olicy **s**imulator **s**ystems* (VDPSS).

3.1.2 Building policy simulator systems: abstraction processes

Abstraction processes transform a messy problematic situation into a structured well-defined problem that allows to design policies to tackle it [115]. Abstraction processes to build policy models create an "artificial world" that serve to support decision-making (especially noticeable in agent-based simulation for policy-making).

These abstraction processes can be illustrated with a simple example. Take, for instance, a urban dweller. As a real person, this citizen is a complex human being "A_C". Nonetheless, considering domestic water use as policy domain, this agent can be reduced to those relevant aspects in the realm of water domain, becoming a simplified water user "$A_\mathcal{W}$". This agent is then modelled: its relevant behaviour is synthesised in some actions (e.g. use water, adopt conservation practices, etc.) and attributes (e.g. environmental values, income, etc.), constituting a citizen-model "$A_\mathcal{I}$". Now, this model is implemented in a simulation platform (i.e. as computer or software agent) to explore some phenomena. Theoretically, the simulated behaviour of the agent "$A_{\widetilde{\mathcal{W}}}$" is similar to that one of the real agent in the domain "$A_\mathcal{W}$".

Thus, **designing policy simulator systems involves multiple abstraction processes** (Assumption 3) (Fig. 3.2)[2]:

(1) First, the "complete" reality "W_C" is reduced to that part of the world that is relevant for the policy issue: "$W_{\mathcal{W}}$". This process defines the policy domain: the purpose of intervening, the stakeholders involved and their behaviour, the dynamics of the system, etc.

(2) Second, the relevant world "$W_{\mathcal{W}}$" is translated into a model of the world "$W_{\mathcal{I}}$". Thus, this process aims to simplify "$W_{\mathcal{W}}$" in order to provide explanations and inferences (See [73]).

(3) Third, this model "$W_{\mathcal{I}}$" will be implemented in a simulation platform (e.g. NetLogo, Repast), and hence expressed in a particular programming language: "$W_{\mathcal{T}}$".

(4) Eventually, the implemented model "$W_{\mathcal{T}}$" is *run* to generate a simulated world "$W_{\widetilde{\mathcal{W}}}$" that emulates "$W_{\mathcal{W}}$".

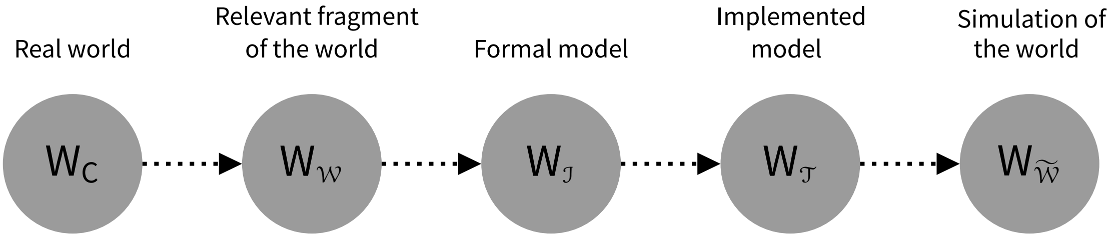


Figure 3.2: Abstraction processes involved in building a model for policy simulator systems

3.1.3 Value-driven policy models

These abstraction processes are not trivial. Complex decision situations (as are policy decisions) are usually addressed by narrowing the mental representation and considering only a limited set of objectives, leaving out other relevant aspects [34]. Indeed, conceptions of what is desirable are usually constrained by what is thought to be possible [250]. Elements that do not exist in the model cannot be affected by any intervention, but this does not mean that they will not be altered in the real world. Consequently, in policy-making, amorality (i.e. absence of values and their meaning in the domain) may lead to policies with detrimental social outcomes; not because there are objective rightful value judgements, but rather because such absence of values may leave essential aspects out of the judgement —consciously or unconsciously.

For instance, de Wildt *et al.* [269] explored how the energy transition involves multiple value conflicts (e.g. *equity vs. efficiency*), and pointed out that unsatisfied expectations concerning values may eventually lead to social acceptance issues, which may jeopardise the transition. For that reason, not having those sensitivities, may hide these value conflicts, resulting in counterproductive policies (see also [44, 105]).

For this reason, it is necessary to develop guidelines to make decision-makers and modellers to be aware of these issues, as they may unconsciously project (only) their own values onto the model. The intuition here is that considering multiple values in modelling can lead to expand the relevant world, producing a cascade effect (because "policy problems" are defined by taking values into account). **This value-driven approach can benefit agent-based modelling for policy-making** (Assumption 4): ethically, because values are taken into account and *imbued* in the model; and scientifically, because it may foster a further exploration of the effects of policies in the form of research questions. A methodological proposal to address this aspect is further explored in Chapter 8.

---

[2] Notice that this process does not conflict with the one presented in Sec. 2.2.2, which presents the main stages of policies' impact assessment. The present process simply details those abstraction processes involved in *designing* the problem to be *solved* with a computer system and agent-based simulation.

3.1.4 Policy paradigms

Interestingly, the abstraction of the world may be restricted by policy paradigms. Paradigms are constituted by dominant worldviews and consolidated values, that are seen as self-justified common sense [49, 181]. Furthermore, they may establish particular measures of success and failure in policy domains. As a consequence, paradigms influence how collective problems are defined and approached, thus shaping policy models.

Since paradigms shape the beliefs of stakeholders about the policy world, those visions are largely shared. Thus, **paradigms are reflected in the beliefs of agents in the policy domain** (Assumption 5). In practical terms, **a paradigm frames the policy problem by establishing a limited set of values that are to be considered in the policy domain (and their understanding) and the set of "reasonable" options to consider** (Assumption 6).

3.1.5 Modelling policy proposals as *policy-schemas*

The *cause-effect* schema is highly present in human cognition [156], and it is the common basis to make sense of the world. Then, the clearest way to model policy proposals is to distinguish between two main components: (a) *policy ends* (i.e. *intended effect*), which are those objectives that are meant to be achieved with the intervention (thus, *better* states of affairs); and (b) *policy means* (i.e. *potential cause*), that are those mechanisms that will change the state of affairs. In fact, a main part of models consists of defining the *causal mechanism* of the domain (that will be simulated to explore the actual effects of a policy enactment).

Typically, policies aim at producing a behavioural change at a social level [247]. Hence, *policy means* aim to produce a change on the relevant world in order to drive the system towards a deemed desirable state of affairs. In practice, they are implemented with diverse *instruments* (e.g. financial, economic, regulatory, informational, etc.); however, these interventions must consider that the agents of the system, for whom the policy proposal is designed, are not passive entities, but seek to further their own interests [236]. In the same vein, the assessment of the degree of achievement of those *policy ends* relies, in practice, on indexes and *indicators* —either quantitative or qualitative— [112], computable from variables of the relevant world. **In essence, values take two relevant implications when talking about public policies: (a) *values as motivators of behaviour*; and (b) *values as social standards*** (Assumption 7).[3]

With this in mind, **a *policy-schema* is the way the policy proposal is meant to be implemented, and is composed of two main constructs** (Assumption 8):

(a) *Policy means*, that aim to produce a social behavioural change in order to drive the system towards a desirable state of affairs, and are implemented with *instruments* that guide the social activity towards the policy ends.

(b) *Policy ends*, that define desirable state of affairs intended to be achieved, and are expressed through *indicators*, which are computed from variables that are observable in the relevant world, and whose actual assessment inform whether policy objectives are achieved or not.

For example, an *end* in urban water policies may be to guarantee a (quantitatively) safe supply of water, which may be represented by a combination of *indicators*, such as the available water (e.g. days of supply for an average person). The *means* of a policy with that *end* may be controlling water demand

---

[3] Sekera [230] recognises the criticism of the "measurement mania", but points out that this is indeed a need in public economy for performance measurement. According to this author, the reasons for constructing a meaningful method for measurement are fourfold: (1) to determine whether an intended state has been achieved; (2) to improve results; (3) to inform decision-makers (who make ongoing decisions about authorisation and funding); and (4) to inform the public.

of users, which can be implemented through *instruments* as higher water fees for large consumers to discourage abnormally high water use.

It is important to have these four elements in mind (i.e. *ends*, *indicators*, *means*, and *instruments*). Taking *indicators* but not *ends* may lead to consider, mistakenly, that such *indicators* are actual policy goals by themselves (e.g. Gross Domestic Product instead of economic development to improve social well-being). Likewise, to have *instruments* without *means* may hide the main reasoning behind the application of those instruments. Finally, having *ends* but not *indicators*, or *means* without *instruments*, may lead to "too ideal" policies that lose specific implementation issues on the field.

3.1.6 Value-driven assessment

In practical terms, values serve as standards to assess policies and states of affairs. However, values may take different meanings depending on the context and domain of interest.

A state of affairs is characterised by the properties that hold that state of affairs (i.e. variables, facts, etc.). These properties are assessed with respect to the values hold dear, to discern whether a state is "good" or not. Furthermore, assessing multiple states (real or hypothetical) allows to generate preferences over the state of the world.

For policies, it is convenient to distinguish between *effective policies* and *good policies* [192]. *Effective policies* are those policies whose outcome is consistent with the policy declarations, in an objective sense. In contrast, *good policies* are those whose outcome improves the state of affairs from a subjective point of view (thus, according to one's own values). For instance, a policy may be *effective* from a policy-makes' point of view, but not *good* for all stakeholders. In this connection, a societal challenge consists of negotiating and agreeing on those consensual social values, in order to develop a system to evaluate whether a policy is *socially good*.

Besides, a policy is *acceptable* when the stakeholders agree that the outcome is sufficiently similar to the outcome of the most *favourable* alternative intervention (including non-acting) and, therefore, they are not likely to act to produce any policy shift [44].

Hence, a policy can be assessed with regard different standards (e.g. *effectiveness*, *desirability*, *acceptability*). Interestingly, **standards for policy assessment involve values** (Assumption 9). For instance, *effectiveness* depends on the definition of the "policy problem" (thus, policy-makers' values), while *acceptability* depends on the values of the stakeholders of the system.

That said, one should notice that taking multiple values into account does not solve ethical dilemmas in policy decisions, basically because these cannot be solved scientifically: there will never be an objective way to establish a correct system to make value judgements. This can only be made politically and on the basis of social consensus of what are considered to be social values. With this in mind, the idea is that policy simulator systems should be broad and transparent enough so as to present value conflicts clearly in order to allow diverse users to be widely informed. In fact, it has been reported that decision-makers, despite holding values and knowledge relevant for particular decision situations, may be unable to generate a comprehensive list of value-based objectives, presenting better performances when external was involved assistance [34, 133] —which reinforces the need to include values in simulator systems for policy-making.

3.2 Working with values in policy simulator systems

To further work on a conceptual framework that considers *values* as a primary entity of policy simulator systems, it is necessary to make explicit assumptions:

3.2.1 Cognitive view of values

Values are cognitive constructs that are involved in the rational behaviour of an agent and serve as standards, refer to desirable goals and transcend specific actions (Assumption 10). In particular, values guide perception, evaluation, and decision-making in any situation:

- Perception: *"What is relevant and deserve attention?"*
- Evaluation: *"Is the state of affairs good enough? If not, should anything be done to address the situation?"*
- Decision-making: *"What are the options? What is the action to be done?"*

Values are present when reasoning within an environment:

- In a *local environment*, individuals can perceive their situation directly, and therefore make choices whose consequences are more easily traceable and computable for them. For instance, they include micro-economic decisions (e.g. what is the cost, the impact on the budget, the utility of the good or service, how it will affect the family well-being, etc.).
- In a *global environment*, reasoning involves socio-political beliefs, but the context is not directly or easily perceivable for all the agents (because the environment involved is more complex). Hence, individuals make choices whose effects on the social outcome and on their own well-being may be unclear.

In both spheres, agents are *rationally bounded*, making use of several cognitive heuristics and being affected by multiple biases [74, 130, 3]: for instance, values are present in agents' political behaviour because they are involved in determining ideological worldviews and trust in political actors.

The key aspect is that **values play two main roles: (i) as individual standards that motivate the conduct of the agents *locally*; and (ii) as social standards that concern the society of agents as a whole** (Assumption 7). An agent can hold both types of values, but the subject, scale, and situation of activation differ:

(i) *Motivational values* refer to those values towards satisfying personal needs and self-esteem. For instance, *achievement*, *power*, *wealth*, or *stimulation*.
(ii) *Social values* concern the public interest, needs, and well-being of the society at large (i.e. supra-individual entities). These values can be projected onto a policy-schema, and are the basis for policy assessments. For instance, *equity*, *privacy*, *innovation*, or *community's wealth*.

As mentioned before, there is no consensual taxonomy of values, since multiple value items can be considered depending on the focus. Hence, there is no commitment with any particular theory. That said, *motivational value items* can be drawn from the Schwartz Theory of Basic Values [221], which is the theory most supported on empirical evidence. In contrast, *social value items* are more disputed, due to their (social) normative load. In Chapter 2 different theories have been mentioned and may be useful for this matter (e.g. Witesman and Walters' Public Service Values [270]).

3.2.2 Consequentialist view of values

Values have to be made observable in a state of affairs in order to conduct policy assessments. In this spirit, a consequentialist view of values is adopted: **the focus of value-driven decisions is on their consequences in the (simulated) social system** (Assumption 11). In other words, it is not the value itself what is under discussion —which involves philosophical stances—, but rather how to interpret it with the outcomes of an action.

Since values are abstract in nature, they need to be translated into operational representations in accordance with the policy domain, the social context, and their social understanding. Values cannot be measured directly —or at least, there is not a known, consensual method to do that. However, for analytical purposes, one can say that some elements of the state of affairs do reflect the value, which can be measured and assessed. Consequently, the effects of an action or policy on those elements that reflect the value can also be measured.

Hence, **values can be projected onto indexes or indicators that stand for those values** (Assumption 12), allowing to make value judgements by assessing these elements. The main advantage is that values (whose meaning is quite elusive) are not discussed in abstract terms, but rather in specific measurable consequences. This assumption entails that only some relevant aspects of the world are taken into account, namely those that are captured by the indicators or indexes, making the choice of indicators an important issue (see Sec. 3.1.2). In Chapter 8 this instantiation process is further illustrated.

According to this view, an action "a" may promote a value "α" and demote a value "β" depending on their consequences on the corresponding indicators (within a particular domain and particular context). An action "a" is aligned with a value "α" if the outcome improves the state of the world with respect to the understanding of value "α". Likewise, one state of the world "σ" is better than another state "τ" with respect to a value "α" when it has a better valuation for those indicators that reflect the value. In this way, one can conclude that an action is better than another with respect to a value by comparing the effects each one would have in the state of the world.

Despite having limitations, the consequentialist view enables a practical analytical approach for computational models. Notice that this view does not assume that the consequences of an action justify the action, but rather it provides a tool to explore the link between how values are understood in a particular context and the actions that are relevant to it. Although a consequential view fits into the research nature of social simulation (since the aim is to discuss about simulated effects on the model variables), an analysis of the consequences through simulation does not eliminate the responsibility of the actual decision-maker. Accordingly, there may be actions that should not be acceptable under any circumstances, although their effects (informed by simulation) are "good". In this connection, models are abstractions, and hence they may be incomplete and biased.

Interestingly, adopting a consequentialist view of values computational models for policy assessment leads to particular issues. One issue is that it is unreasonable to consider that all the consequences of all possible actions can be taken into account. For this reason, consequentalism involves necessarily the choice of a scope and scale of analysis when assessing alternative actions: namely, it requires to decide to what extent the consequences are going to be considered (i.e. spatially, temporal, socially, etc.). In fact, restricting the scope and scale of assessment is used to establish the horizon beyond which agents are assumed to not have moral responsibility (see [15]). Another issue has to do with having an inappropriate ontology, namely that indicators are missing or wrong and thus the values is not appropriately represented in the model.

3.2.3 Commensurable view of values

Another working assumption has to do with the way several values are simultaneously considered when assessing something. Although one may argue that values are incommensurable (that is, they cannot be measured on a common scale or compared in an objective sense) [116, 195], the fact that individuals solve value conflicts in their everyday lives cannot be ignored.

Human beings are known to use unspecified and inconsistent aggregation models in their everyday lives, besides being subject to framing effects and other cognitive biases. Hence, value conflicts may be solved *ad hoc* by invoking the relative importance between values and some form of satisficing combinations [235]. In this sense, individuals do not choose the best solution, but rather acceptable solutions, by using simple heuristics.

In extremely simple terms, this means that, even if there is a dilemma or value conflict, an individual ends up making a decision. Hence, values are, at least, *cognitively commensurable* for any agent (because choices that involve multiple values are eventually made). Evaluation and aggregation mechanisms must serve an agent to discern whether a state of affairs is better or worse than another one with respect to multiple relevant values, and consequently, to discern whether an action or policy is good or not.

This view is also true for policy decisions, which involve multiple conflicting values, leading to complex choices to make. For example, an agent involved in a decision situation that presents a conflict between *safety* and *privacy* will, eventually, make a choice —often called a 'tragic choice', because, objectively, the gain in one value cannot cancel the loss in the other value [116].

In fact, in order to assess alternative options in a methodical way, it is not unusual to assume that values are commensurable [195, 197]. For instance, the value of *environmental sustainability* may be measured and evaluated in a different way that *economic profit* is, but evaluations with regard to these two values can eventually be aggregated in order to compare different policies. Methodologies like Cost-Benefit Analysis (CBA) or Multi-criteria Assessment (MCA) are used to assess alternative policies by enabling the evaluation and aggregation of multiple values [171]. Other sophisticated approaches have been elaborated to reframe the MCA in order to increase their applicability. For instance, the Social Multi-criteria Evaluation (SMCE) aims at considering the preferences of the multiple stakeholders involved for assessing options [171]. Also, the Integrated Value Model for Sustainability Assessment (MIVES), developed for evaluating alternatives in engineering taking environmental and social concerns into account [198], establishes a systematic protocol to generate aggregation models based on (a) *value functions* that score alternatives with respect to chosen value-representative indicators (see [9]) and (b) *value hierarchies* elicited through the Analytic Hierarchy Process (see [213]).

In any case, **when multiple values are involved in decision-situations, values can be made *commensurable* (hence, they can be aggregated and compared)** (Assumption 13). The system to do so is referred to as "*value assessment frameworks*". This does not imply necessarily that these systems are simply functions that aggregate multiple values and return a score that represents the overall *utility* [1]. Rather, it means that agents are able to consider multiple values and solve value trade-offs in a particular situation, and eventually make a decision or perform an action. The evaluation and aggregation mechanisms used in this system may be diverse in their sophistication.

Value assessment framework

In precise terms, a *value assessment framework* affords to evaluate objects and aggregate multiple relevant values, allowing to make decisions and perform actions. In this view, a *value assessment framework* is constituted, at least, by three fundamental components:

(i) the *value system*, that provides the values and their understandings (especially, the indicators that stand for those values);
(ii) the *evaluation mechanisms*, that provide the actual procedure to assess an object with regard to the relevant values and their indicators (e.g. value functions [9], satisficing thresholds, normalisation methods, etc.); and
(iii) the *aggregation mechanisms*, that enable to aggregate multiple valuations (e.g. weighted mean of scores, hierarchies and combinations of satisficing thresholds, etc.).

Chapter 4

Conceptual framework

The purpose of this chapter is to provide a conceptual framework to characterise the ***v**alue-**d**riven* ***p**olicy **s**imulator **s**ystems* (VDPSS). It aims at supporting the creation and use of policy simulator systems.

The main assumption of this conceptual framework is that policy-making is a value-driven activity, involving two complementary perspectives of values: first, those social values that concern the society at large; and second, those values that motivate the stakeholders of the system. In simple terms, policy-making involves coordinated collective activities whose centrepiece is that policy-makers design and enact a policy to govern the activity of some targeted stakeholders in order to reach a *better* state of affairs with respect to some values within some particular domain.

The use of policy simulator systems aims at providing support to improve or preserve particular states of affairs (and not only understand the dynamics of a societal system), which involves values and ethical views of the users and designers. For this reason, it is convenient to develop a conceptual framework that contextualises simulation for policy-making as a value-driven activity, and presents concepts to support the building of models for policy simulation.

Naturally, policy-making in real life is much more complex than the representation in this conceptual framework. Nonetheless, the framework provides the essential elements to build policy simulator systems by establishing an objective frame of reference where "policy assessment problems" may be worked out to produce insights that support real-world policy-making.

4.1 Outline of public policy models

In loose terms, a public policy is an intervention "Π" to improve a *fragment of the real world* "$\mathcal{W}_{\mathfrak{D}}$", deemed relevant for addressing a collective problem in a particular domain. The centrepiece of this policy intervention consists of two elements: (i) the *policy ends*, that set the goals of the intervention, whose degree of achievement will be used to assess the effectiveness of the policy; and (ii) the *policy means*, that set the instruments to drive the societal system towards those ends. The *causal mechanisms* of the societal system will transform that "impact" produced by the intervention (i.e. *policy means*) into "effects" (i.e. *policy ends*).

Preliminary versions of this chapter were presented in the First International Workshop on Socio-cognitive systems (SCS18) (see [188]), the 14th Annual Conference of the European Social Simulation Association (SSC18) (see [187]) and the 18th Mexican International Conference on Artificial Intelligence (MICAI19) (see [186]).

Since the societal system is complex, its dynamics are not always well-known or fully predictable, and consequently the policy outcome may differ from what has been intended. This is one of the several reasons that justify the computational simulation of public policies: not to predict the actual outcome in the real world, but to anticipate relevant information accurately [6]. Given this purpose, it is necessary to abstract those *causal mechanisms* from the *fragment of the real world* "$\mathcal{W}_{\mathcal{D}}$" in order to simulate faithfully the societal system. In other words, the elements, laws, and dynamics of the *fragment of the real world* "$\mathcal{W}_{\mathcal{D}}$" are represented into a model of this domain —hereinafter termed as *policy domain* "$\mathcal{D}$".

4.1.1 Overview of the basic components

A model of a public policy includes, at least, two more components besides the representation of the real-world intervention (i.e. *policy-schema* "π") and the representation of the domain ("$\mathcal{D}$"). First, public policies concern multiple *stakeholders* that act in the domain and are involved in the collective problem. In practical terms, different stakeholders behave differently in a social space in order to further their goals, because they hold different values and interests, and hence have particular understandings about how to act and how to judge the outcomes in diverse situations.

Second, any public policy involves *values*, since the intervention aims at improving the state of affairs —that is, promoting certain states of the world over others. Values provide the reasons that justify the intervention, because they define what is *desirable* in a "policy problem". Furthermore, the other stakeholders (aside from the policy-maker) that act in the domain also behave in accordance with their values and interests. Therefore, and expressed in practical terms, values determine the standards that serve to assess the state of affairs.

In sum, a *public policy model* contextualises interventions in a domain of interest by taking values, stakeholders and causal mechanisms into consideration, thereby allowing to reason accurately about the policies and their social effects. In more precise terms, a *public policy model* "$\mathcal{M}$" consists of four main components (WDef. 1):

(i) A *policy domain* "$\mathcal{D}$", which is the representation of the fragment of the real world that the policy is intended to improve; namely, the elements, laws, and causal mechanisms that describe the behaviour of the societal system.
(ii) A *population of stakeholders* "$\mathcal{A}$", socially interconnected, who participate and interact in the domain to further their interests according to their socio-cognitive decision-making models.
(iii) A *value model "$\mathcal{V}$", that establish the relevant values in the domain, from a twofold perspective: (i) those that serve to design the intervention and assess the social outcomes; and (ii) those that are part of the socio-cognitive models of stakeholders.*
(iv) A *policy-schema* "π", which is the intervention itself; namely, the *policy ends* "π_e" (i.e. what is to be improved) and the *policy means* "π_m" (i.e. how to achieve that improvement).

Working definition 1 *A **value-driven public policy model** "$\mathcal{M}$" is an abstraction of a policy domain "$\mathcal{D}$", a population of stakeholders "$\mathcal{A}$", a value model "$\mathcal{V}$", and a policy-schema "π" (Expr. 4.1).*

$$\mathcal{M} = \langle \mathcal{D}, \mathcal{A}, \mathcal{V}, \pi \rangle \tag{4.1}$$

4.1.2 Computational simulation of public policy models

Public policy models can be simulated in a computational system in order to provide practical insights and feedback to support policy-making. The simulated outcome of any intervention is unequivocal and coherent with how the model has been defined (without entering the verification and validation of the model itself). Hence, the simulation of public policy models may be useful for reasoning on more complicated systems than those that are mentally tractable, contributing to avoid intuitive errors.

In this spirit, a formal *public policy model* “$\mathcal{M}$” is *implemented* in a *simulation platform* “$\mathcal{T}$” (e.g. NetLogo, Repast, MASON, etc.), which includes some specification conventions that serve for coding the formal model into a computational model (for instance, a programming language based on particular commands and syntax rules).

For the sake of conducting simulation experiments, one should distinguish between control variables (associated to independent variables) and output variables (associated to dependent variables). Thus, the formal model “$\mathcal{M}$” may have multiple *input variables* that are “open” and require to be instantiated before the simulation, and multiple *output variables* that require to be selected to configure the simulation output.

It is useful to consider the following insights: first, it is important to distinguish between those *input variables* that correspond to elements that are subject to be changed in the real world by means of public policies, and those parameters involved in the causal mechanisms that underlie the system behaviour. Second, there may be *output variables* that are used as performance indicators to assess the outcome of the policy enactment, and others that, despite not being part of the intervention itself, are monitored for better examination of the system behaviour. In other words, only some *input variables* and *output variables* are associated to the policy intervention (i.e. *policy-schema* π) (Fig. 4.1). For the sake of example:

- The *policy domain* “$\mathcal{D}$” may have unspecified parameters that regulate the *causal mechanisms*, which may be modelling constructs, variables that are empirically not well-known, or variables that are not alterable by a policy intervention in the context of study (e.g. social influence mechanisms, event likelihood, etc.).
- The socio-cognitive models of the *stakeholders* “$\mathcal{A}$” may rely on adjustable parameters that control their decision-making (e.g. risk aversion, etc.).
- The mechanisms used in the *value model* “$\mathcal{V}$” may set some parameters that are open (e.g. tolerance thresholds, aggregation rules, etc.).
- The *policy-schema* may adjust the exact value for an *instruments* (e.g. a subvention for technology adoption is set to 5 %) and chooses the *indicators* to assess its performance with respect to a *value* to be imbued (e.g. adoption rate to assess ‘innovation’).

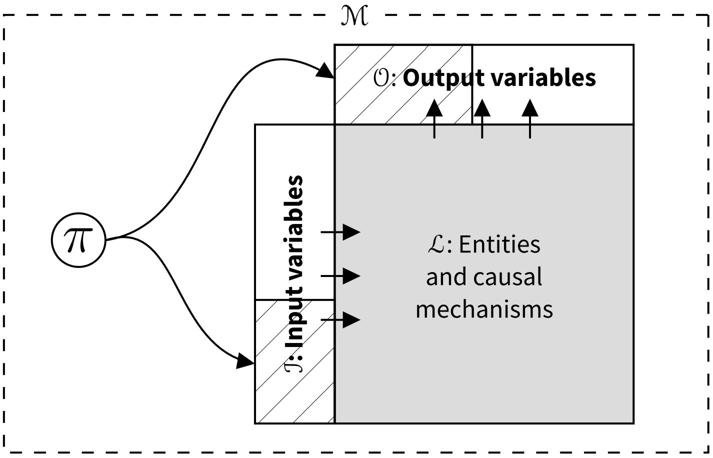
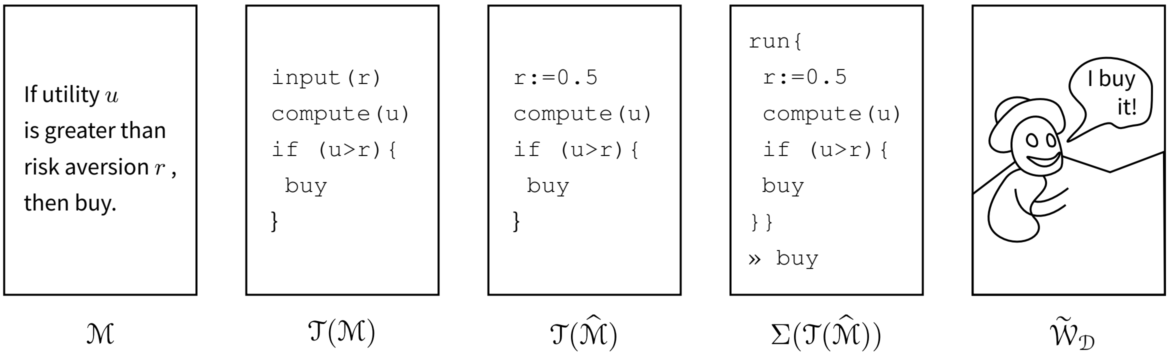


Figure 4.1: A view of the public policy model. The formal model “$\mathcal{M}$” may have some “open” variables that have to be instantiated for the simulation. In particular, the *policy-schema* “π” sets the *instruments* and the *indicators* of the intervention.

The formal model "$\mathcal{M}$" must be "closed" before its simulation: all the variables must be instantiated, thus constituting an *instantiated model* "$\widehat{\mathcal{M}}$". Commonly, the model is instantiated through a user interface. Then, the computational model is then executed ("Σ") to generate a simulated world that emulates the *fragment of the real world* "$\mathcal{W}_{\mathfrak{D}}$" (creating a *simulated reality* "$\widetilde{\mathcal{W}}_{\mathfrak{D}}$" that is similar to the real world "$\mathcal{W}_{\mathfrak{D}}$") (Fig. 4.2). Interestingly, there is an interdependence between the real world and the simulated world: first, "$\mathcal{W}_{\mathfrak{D}}$" shapes "$\widetilde{\mathcal{W}}_{\mathfrak{D}}$", because one is an image of the other; and second, "$\widetilde{\mathcal{W}}_{\mathfrak{D}}$" may affect "$\mathcal{W}_{\mathfrak{D}}$", because it may lead to interventions in the real-world, hence transcending the simulation.[1] [2]



If utility u is greater than risk aversion r , then buy.

```
input(r)
compute(u)
if (u>r){
 buy
}
```

```
r:=0.5
compute(u)
if (u>r){
 buy
}
```

```
run{
 r:=0.5
 compute(u)
 if (u>r){
 buy
}}
» buy
```

Figure 4.2: Illustration of the generation of a *simulated reality* by means of the implementation and simulation of a model

4.1.3 Simulation environment

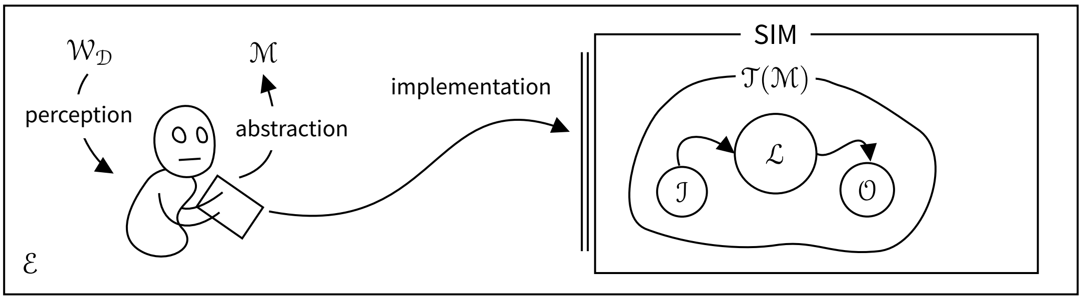
The *simulation environment* "$\mathcal{E}$" consists of the elements that anchor the simulation process into the real-world. These elements may be physical (e.g. users, experts, simulator system, equipment, etc.), but also epistemological (e.g. design of experiments) and contextual (e.g. stage of the policy cycle, use of the simulation, etc.). The *simulation environment* "$\mathcal{E}$" is also a socio-cognitive technical system, that includes, among other elements, a *simulator system* "SIM", the user, and the *experimental setting* "χ" (Fig. 4.3).

The *simulator system* "SIM" is an artifact that can execute (an instantiation of) a computational model of public policy to generate a *simulated reality* "$\widetilde{\mathcal{W}}_{\mathfrak{D}}$" (WDefs. 2 and 3). There is an interface between a *simulated reality* —created by the simulator system— and the real world —that contains the simulator system. The user of the simulator system observes the *simulated reality* through the interface.

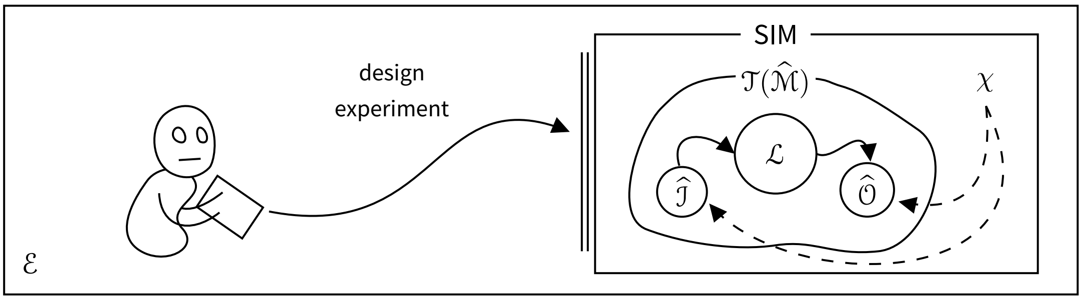
---

[1] The *simulated reality* "$\widetilde{\mathcal{W}}_{\mathfrak{D}}$" does not exist in factual terms. Nonetheless, this notion may help to imagine that an "artificial reality" is created to emulate the real world. There is multiple cultural references that can be invoked to illustrate the notion of *simulated reality*. For instance, in the world of cinema, one of the most iconic movies that shows such concept is *The Matrix*. It depicts a dystopian future in which human beings are trapped inside a *simulated reality*, which they believe to be the reality because they ignore that there is a "real world" outside the simulation. In the world of video-games, there is plenty of examples —practically, any simulation game. One illustrative example is *The Sims*, a game where the player creates virtual human beings called "sims", who live in houses and need some help to satisfy their needs and desires. In factual terms, these sims are nothing but a software program. However, from a metaphysical view, they live inside a *simulated reality*, which they "think" is the real world.

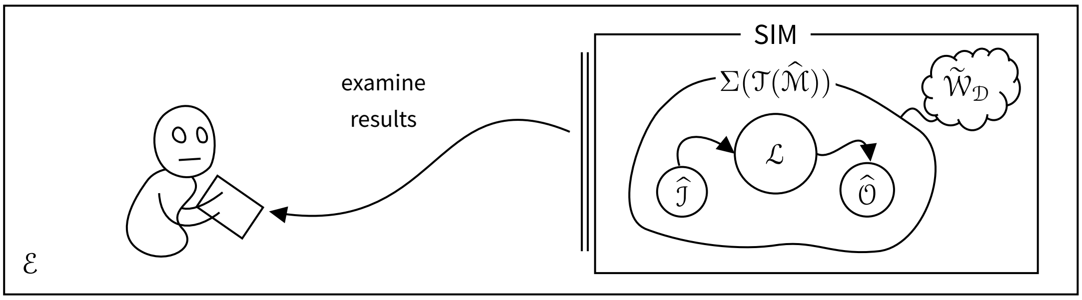
[2] A model "$\mathcal{M}$" may include random variables. In this case, "$\mathcal{M}$" considers such variables as a random distribution. The execution of the instantiated model "$\widehat{\mathcal{M}}_j$" leads to different results due to the randomness. The exact value for a random variable is the result of the implementation "$\mathcal{T}$" (the algorithms for generating random numbers) and the execution "Σ" (the exact number produced by the algorithm and a seed) of the random distribution. As a consequence, it is common to make a statistical process of the simulation outputs of the same model "$\widehat{\mathcal{M}}_j$" (for instance, the average of an output from multiple runs —i.e. "$\overline{\Sigma}$").



(a) The modeller (in the *simulation environment* “$\mathcal{E}$”) abstracts the relevant *fragment of the real world* “$\mathcal{W}_{\mathcal{D}}$” concerning a domain into a *value-driven public policy model* “$\mathcal{M}$”, that is *implemented* “$\mathcal{T}(\mathcal{M})$” to build a *simulator* “SIM”. The *simulator* “SIM” presets *input variables* “$\mathcal{I}$” and *output variables* “$\mathcal{O}$”, which will be specified later to conduct different experiments. Different preselections of *input variables* and *output variables* correspond to different *implementations* and, therefore, different *simulators*.



(b) The modeller programs an *experimental setting* “χ” using the *simulator* “SIM” (for instance, to test the effect of different values for an *input variable* on a particular *output variable*). This instantiates the *input variables* “$\widehat{\mathcal{I}}$” and selects the *output variables* “$\widehat{\mathcal{O}}$”, thereby constituting an *implementation* of the *instantiated model* “$\mathcal{T}(\widehat{\mathcal{M}})$”.



(c) The modeller *runs* the *implemented instantiated model* “$\Sigma(\mathcal{T}(\widehat{\mathcal{M}}))$” in the *simulator* “SIM” to examine and analyse the results. The execution of one instantiation of the model generates a *simulated reality* “$\widetilde{\mathcal{W}}_{\mathcal{D}}$” —which emulates the behaviour of the *fragment of the real world* “$\mathcal{W}_{\mathcal{D}}$” under given circumstances.

Figure 4.3: Illustration of the main processes involved in building and using a simulator system in a simulation environment (leaving model verification and model validation out of the illustration).

Working definition 2 *A **simulator system** “SIM” is an artifact, characterised by a platform “$\mathcal{T}$”, that can conduct simulations “Σ”, whose core is the implementation of a public policy model “$\mathcal{T}(\mathcal{M})$”, and hence is able to simulate different instantiations of that public policy model “$\Sigma(\mathcal{T}(\widehat{\mathcal{M}}))$”.*

Working definition 3 *Given an instantiation of the implemented public policy model “$\mathcal{T}(\widehat{\mathcal{M}})$”, a **simulation** “$\Sigma$” is the execution of the implementation to generate a simulated reality “$\widetilde{\mathcal{W}}_{\mathcal{D}}$”, which is conducted by a simulator system “SIM”.*

The knowledge from using a simulator is produced usually by analysing multiple *simulated realities*, which correspond to different instantiations of the model. This is established by the *experimental setting* χ, that configures the experiments and protocols, the treatment and analysis of the simulation outputs, and the visualisation and communication of the results (WDef. 4). This allows to study policies with regard to standards that are difficult to be explored in the real world, because the conditions of the real system are never the same (for instance, statistical analysis of system sensitivity, robustness, etc.).

Working definition 4 *An* ***experimental setting*** *"χ" is the set of experiments, protocols, and techniques that define how the simulator system SIM is going to be used to generate knowledge about the real world "$\mathcal{W}_{\mathcal{D}}$" from analysing one or multiple simulated realities.*

Considering that the *simulator system* "SIM" is used to support policy-making, the user is assumed to hold the role of *policy-maker* of the *simulated reality* (regardless of whether it is an actual policy-maker, a developer, or a consultant team). The user is able to enact interventions in the *simulated reality* and draw conclusions to support policy decisions in the real world, based on the evaluation of the *simulated reality* according to its own value system (see Sec. 4.4.1). Presumably, these values are aligned with the responsibility of the role of *policy-maker*, and many may be present in assessing the policy despite not having a direct representation in the *simulated reality* (e.g. transparency, responsiveness, accountability, compliance, etc.) (see Chapter 6).[3]

4.2 Policy domain

The *policy domain* refers to the representation of the fragment of the real world relevant for addressing a collective problem in a specific domain.

The (formal) representation of the *policy domain* is the model that recognises entities, relationships, events and actions that take place in the domain of interest. In basic terms, it defines the causal mechanisms and laws that apply to the model of the societal system. Thus, this model (i.e. *policy domain*) prescribes a specific *ontology*, which is the formal account of the entities considered to be involved in some system, and the relationships between them [96]. For instance, in the domain of urban water management, the ontology may involve entities such as water, households and water fees, and the relationships express that households are supplied with water, and that households pay proportionally to the water fees and their actual water use.

Working definition 5 *The* ***policy domain*** *"$\mathcal{D}$" establishes the entities and interactions that are involved in some policy domain, prescribing the laws and causal mechanisms "$\mathcal{L}$" and defining the state of the world "$\mathcal{S}$".*

The ontology establishes the primitive entities that define the *state of the world* as well as the actions and events that can change it. A *state of the world* "$\mathcal{S}$" is an instantiation of the variables that constitute it; it is objective and unique at any moment in time and it is shared by all the participating agents (WDef. 6).

---

[3] The modeller, the developer, and the user may not be the same agents (see [87]). Hence, the *public policy model* "$\mathcal{M}$" is *compatible* with the values of the user as long as the user is able to take these values can be reflected in the simulation output (because indicators that reflect those values have been included in the model). For the sake of simplicity, in this chapter is assumed that the modeller, the developer, and the user are the same agent, and therefore the model is perfectly *compatible*.

Working definition 6 *Given an ontology that involves a set of "r" variables $\{\alpha_1, \dots, \alpha_r\}$ which can take values in the spaces $\{d_1, \dots, d_r\}$ respectively, a* ***state of the world*** *"$\mathcal{S}$" is an instance of the set of variables $\{\hat{\alpha_1}, \dots, \hat{\alpha_r}\}$, being $\hat{\alpha_j} \in d_j$. Accordingly, the set of possible states of the world is given by $\{\mathcal{S}\} = \times_{j=1}^{r} d_j$.*

In a historical sense, there is a *state of the world* for each moment in time (that is, the *state of the world* is a "snapshot" of the state of the system that may change each time step of the simulation) (Expr. 4.2). Some policy assessments may rely only on the *state of the world* at a particular time (for instance, at the end of a particular period of time), but others may consider the evolution.

$$\begin{array}{c|cccc|l}
 & \alpha_1 & \alpha_2 & \dots & \alpha_r & \\
t_1 & \hat{\alpha_1}^{(t_1)} & \hat{\alpha_2}^{(t_1)} & \dots & \hat{\alpha_r}^{(t_1)} & \to \mathcal{S}_1 \\
t_2 & \hat{\alpha_1}^{(t_2)} & \hat{\alpha_2}^{(t_2)} & \dots & \hat{\alpha_r}^{(t_2)} & \to \mathcal{S}_2 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
t_\Omega & \hat{\alpha_1}^{(t_\Omega)} & \hat{\alpha_2}^{(t_\Omega)} & \dots & \hat{\alpha_r}^{(t_\Omega)} & \to \mathcal{S}_\Omega
\end{array} \tag{4.2}$$

Since the *policy domain* is a representation of a real-world system, it may be divided into multiple sub-domains (i.e. biophysical, economic, regulatory, technological, social, etc.), which may have different specification languages. This division may allow to further sophisticate the behavioural models of agents with regard to the components of those sub-domains (for instance, an agent may obey the legislation in all cases, but may question the social norms when these jeopardise its goals).[4]

4.3 Stakeholders

Stakeholders are agents —who may be individuals, groups of individuals, or organisations— who are capable of performing some actions in the policy domain (WDef. 7). An stakeholder is characterised, generally speaking, by its role(s) (which determines the affordances in the policy domain) and its socio-cognitive model. The socio-cognitive model involves multiple constructs and factors (e.g. values, beliefs, personality, etc.) that guide their cognitive functions (e.g. perception, evaluation, decision-making). Besides, as stakeholders are assumed to be social agents, they are "interconnected" through a social network (topology, interaction rules, etc.). For the sake of example, a household in the urban water domain is able to use water, influence and be influenced by neighbours, and advocate for a change of tariffs, among many other actions.[5]

Working definition 7 *The population of* ***stakeholders*** *"$\mathcal{A}$" is the set of interconnected agents that participate in a policy domain. A* ***stakeholder*** *"$\mathcal{A}_j$" is an agent capable of performing some actions to change the state of the world, who holds a role(s), and who has a characteristic socio-cognitive model.*

---

[4] *Specification languages* serve to talk about entities and relationships at the formal level (for example, a language that refers to norms (e.g. modal logic), another that refers to actions, another to refer to rhetorical messages, etc.). When the formal model is implemented in the simulation platform, these languages may be mixed up by the programming language itself. For this reason, there is no guarantee that everything expressed in the formal model will be perfectly mapped on the implementation (i.e. not everything formalised is expressible in terms of code so that all the concepts of the model are faithfully implemented).

[5] *Agents* are entities that can perceive their environment and act in consequence according to some internal model of rationality [212]. In more precise terms, agents are entities that are *situated* in some environment; are capable of some *autonomous* action; can process and exchange information (thus, they may have *social abilities* to interact with other agents); are *reactive* to the environment changes and *pro-active* to engage a goal-directed behaviour; may have *heterogeneous* properties; and are *interdependent* with other agents and entities of the environment [273, 200].

4.3.1 Stakeholders as artificial agents

The conceptual framework understands that *policy-makers*, who are authorised to make decisions on behalf of the society, design and institute public policies to regulate a social system in order to change the behaviour of targeted stakeholders and achieve a better state of affairs within some domain of activity (Fig. 4.4). The values of policy-makers (at least, those social values for whose advocacy they have been elected) are involved in the choice of policy means and policy ends.

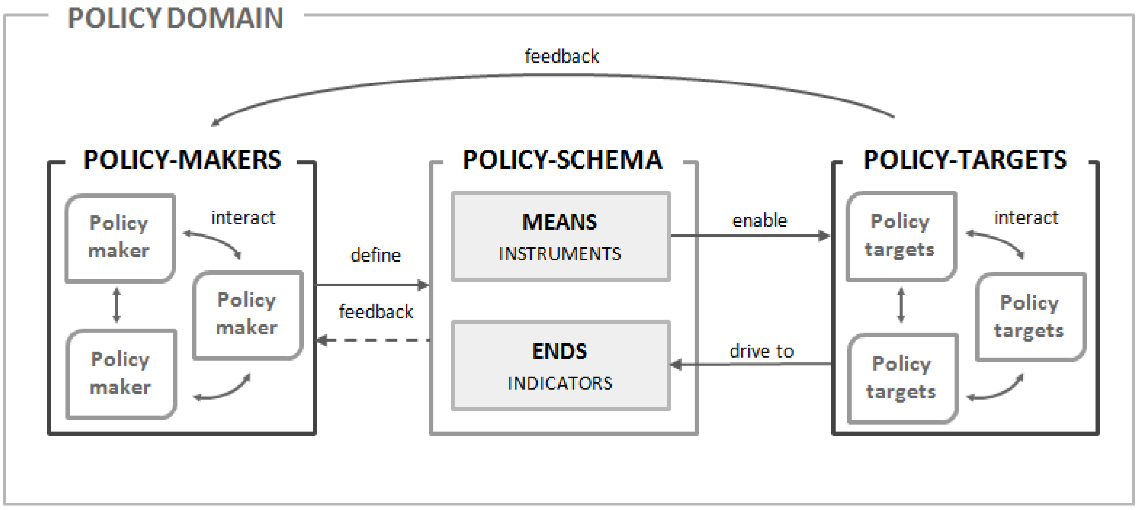


Figure 4.4: Policy-making as a social coordination process. *Policy-makers*, who aim at improving the state of the affairs, institute means and ends in order to govern the activity of *policy-targets*, who are those that are affected by the policy, and whose behavioural change is going to drive the system towards that desirable state. *Policy-makers* monitor the social outcome to re-design, maintain, or terminate the policy. This process happens in a particular policy domain, that contains the relevant part of the world (e.g. social, environmental, economic).

For the sake of simulation, users design and enact a *policy-schema* to govern the activity of *policy-targets* (i.e. targeted stakeholders) in order to reach a desired state of the world. In this case, *policy-targets* are artificial agents, (i.e. computer programs) that exhibit a *rational* behaviour within the policy domain. The focus of the simulation is the behaviour of *policy-targets* as a reaction of the enactment of different policies (Fig. 4.5).

With respect to policy simulator systems, two major roles are assumed: *policy-makers*, that use the tool; and *policy-targets*, that are within the simulation. Within the simulation, the roles of agents in the social system depend on the domain being modelled. For instance, in the case of the urban water domain, the roles may include households, service businesses, and urban industries. These different roles are associated with different affordances and attributes. Back to the previous example, a household may have, among its attributes, the *number of people* in the household (which determines its water use volume) and among its action repertoire: *wash the laundry*, *do the dishes*, and *pay the bills*. In contrast, an industry have other attributes and other actions, which means a different type of participation in the system. Furthermore, among these roles, agents may behave differently if they hold different values.

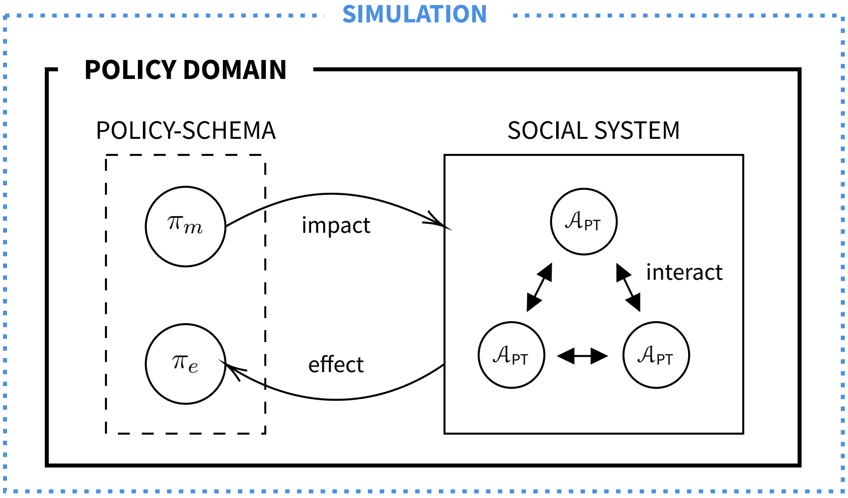


Figure 4.5: Simulation of a policy enactment. Policy means “π_m” are enacted in a social system of artificial *policy-targets* “$\mathcal{A}_{\text{PT}}$”, who adapt their behaviour accordingly and change the state of the world —presumably, towards a desired state of the world defined by the *policy ends* “π_e”.

4.3.2 Basic affordances depending on the scales of concern

Two types of perception scales are distinguished for agents:[6]

(p.1) ***Micro-level* perception**, of those variables that concern the (observable) local environment of the agent, thus being part of its decision-making process. For instance, taking a household as an agent in the urban water domain, the variables of its local environment may include its own income, its water needs or the water fees.

(p.2) ***Macro-level* perception**, of those variables that concern the social system as a whole. For instance, the same household may not be able to observe variables like the total population of the municipality, the average income of its neighbourhood, or the trends in the water use of multiple households (but other agents may be aware).

With this in mind, two types of evaluation scales are differentiated:

(e.1) ***Micro-level* evaluation**, which would correspond to the assessment of the state of the local environment.

(e.2) ***Macro-level* evaluation**, which would correspond to the evaluation of the system at a social level.

Generally speaking, *policy-targets* (e.g. individuals, companies or collectives) focus on its local environment in order to operate, but they can also take insights of the system (for instance, when deciding over long-term actions by considering macro-economic trends). Nonetheless, if *policy-targets* are individual persons, they often have a limited perception of the system as a whole (see Chapter 5).

Notice the distinction between the actual state of the world and that one perceived of agents (i.e. beliefs). The ontology of the model defines what constitutes the state of the world objectively, which is shared by all agents. However, the perception of agents is *partial* (not all the variables of the world-state are relevant or observable); *focused* (restricted to the scale of interest, either micro or macro); *biased* (completed with other types of knowledge, not necessarily factual); and *subjective* (agents may project their values onto different indicators to eventually assess the state of the environment).[7]

---

[6] See Chapter 5 of this distinction being used to model policy shifts within simulations.

[7] An illustrative example of this can be viewed in the work of Sprong *et al.* [242]. They explored the empirical correlation

4.4 Value model

A *value model* "$\mathcal{V}$" establishes the relevant values (i.e. standards of assessment) and their understandings in the public policy model from a threefold perspective: (i) those values that serve to assess the effects of the intervention in the domain; (ii) those values that are part of the rationality of stakeholders; and (iii) those values that serve to assess some additional features of the intervention. In practical terms, a *value model* is based on constructs called *value assessment frameworks*, that define the interpretation and evaluation mechanisms of those held-values (Wdef. 8).

Working definition 8 *The* ***value model*** *"$\mathcal{V}$" is the set of all value assessment frameworks "$\mathbb{V}$": those that serve to assess the effects of the policy intervention in the domain, those that are held by the involved stakeholders, and those that serve to assess some aspects of the intervention.*

4.4.1 Value assessment frameworks

A *value assessment framework* "$\mathbb{V}$" defines the practical way the *state of the world* is assessed with respect to some values that are deemed relevant. It constitutes an explicit expression of how values are made operational, by translating values into factual terms and providing evaluation processes (Wdef. 9).

Working definition 9 *A* ***value assessment framework*** *"$\mathbb{V}$" is a system that allows to assess states of the world with regard to multiple values that are deemed relevant. It is an assemblage of a set of values "V", their pertinent set of factual indicators "i_V", a set of a set of evaluation mechanisms "ϕ_V", and a set of aggregation mechanisms "g" (Expr. 4.3):*

$$\mathbb{V} = \langle V, i_V, \phi_V, g \rangle \tag{4.3}$$

More precisely, for each (i) *value* "V" in the value model "$\mathbb{V}$", there are (ii) a *set of factual indicators* "i_V" onto which the value is projected; and (iii) a *set of evaluation mechanisms* "ϕ_V" for assessing its status with respect to the values through these indicators. Eventually, (iv) a *set of aggregation mechanisms* "g" allows to aggregate multiple value assessments into one overall assessment (Expr. 4.4) (Wdefs. 10, 11 and 12).

$$\left.\begin{array}{ccccc} V_1 \rightarrow & i_{V_1} & \rightarrow & \phi_{V_1}(i_{V_1}) \\ V_2 \rightarrow & i_{V_2} & \rightarrow & \phi_{V_2}(i_{V_2}) \\ \vdots & \vdots & & \vdots \\ V_n \rightarrow & i_{V_n} & \rightarrow & \phi_{V_n}(i_{V_n}) \end{array}\right\rangle g(\phi_{V_1}, \ldots, \phi_{V_n}) \tag{4.4}$$

Working definition 10 *Let a state of the world "$\mathcal{S}$" be defined by the variables $\{\alpha_1, \ldots, \alpha_r\}$. A* ***factual indicator*** *"i_{V_k}" is a combination of a subset of variables $\{\alpha_i, \ldots, \alpha_j\}$ whose purpose is to reflect a value "V_k".*

Working definition 11 *An* ***evaluation mechanism*** *"ϕ_j" is a relation that enables to map the instance of a factual indicator onto a scoring scale (i.e. valuation), thus allowing to compare and elicit preferences between two different instances.*

---

of the economic inequality of the system, measured with the Gini coefficient (i.e. a *factual indicator* to reflect *equity* on the *world-state*), and the individuals' subjectively perceived economic inequality (i.e. the individuals' beliefs about *equity*).

Working definition 12 *An* ***aggregation mechanism*** *"g_j" is a relation that enables to map multiple valuations with regard to different values into onto an scoring scale (i.e. overall valuation), thus allowing to compare and elicit preferences between two different states of the world (with respect to multiple values).*

Example of evaluation mechanism: value functions

Value functions are an example of *evaluation mechanism* that can be used to assess the state of the world (see [9, 198]).[8] In this case, the domain of these functions would be the *factual indicators* that reflect a particular value in that policy domain (i.e. observable variables). Their image ranges from 0 (i.e. lowest valuation score) to 1 (i.e. highest valuation score), which represents the score of those indicators with regard to a particular value. For instance, to assess the state of the world with regard to *environmental protection*, one may use the average households' water use (measured in litres per person per day) as *factual indicator*, and a *value function* that gives the highest valuation score to a water use below or equal to 100 L/p·d, and the lowest valuation score to a water use greater or equal to 300 L/p·d (Fig. 4.6). The scorings with respect to multiple values could be then aggregated using, for instance, a weighted arithmetic mean (i.e. *aggregation mechanism*).

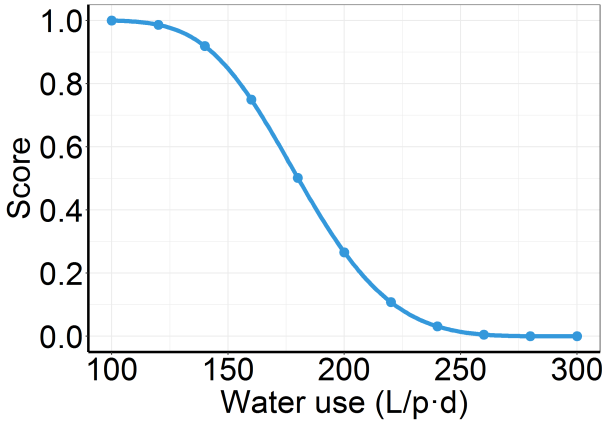


Figure 4.6: Example of *value function* as *evaluation mechanism* for assessing the state of the world with respect to the *environmental protection*

4.4.2 Types of value assessment frameworks

Values and simulated reality

Values must be able to be reflected in the *simulated reality* for two reasons. First, artificial agents may base their decision-making models on value-based reasoning (thus they must be afforded to recognise values and indicators). Second, the user of the simulator will assess the *simulated reality*.

Values in the *simulated reality* are projected onto variables in the *policy domain*. In other words, values have a specific meaning that depends on the way the policy domain has been modelled. The assessment with respect to a particular value is only possible when the model includes variables that correspond to those values. For the sake of example, a policy cannot be assessed with regard to *environment protection* if there are no variables that can reflect those characteristic expectations of the value (e.g. concentration of pollutants, emissions of GHG, etc.). Chapter 8 provides a detailed illustration of how abstract values can be brought into computational models for policy-making.

---

[8] See [9] for further details about how to build *value functions*.

Values of the stakeholders

Each stakeholder has its own *value assessment framework* that serves to resolve trade-offs in situations where several values are simultaneously involved. In precise terms, an stakeholder “$\mathcal{A}$” has its own *value assessment framework* “$\mathbb{V}_{\mathcal{A}}$” that uses to evaluate the worthiness of objects (e.g. state of affairs, policies, etc.) with respect to its held values, and thus provide feedback to its decision-making model. This system represents the *value profile* of stakeholders.

Assessment of policies

Assuming that the simulator is being used by a *policy-maker*, policies may be assessed with regard to *effectiveness*, *acceptability*, and (its own understanding of) *desirability*. These can be viewed as *meta-values* because their actual definition depends on how the domain values in the *simulated reality* are defined. The result of the trade-off between these *meta-values* determines whether to further the intervention (leaving model validation and epistemological issues out when discussing the policy intervention in the real-world). The assessment with regard to these *meta-values* corresponds to different *value assessment frameworks*:

- **Effectiveness.** Policies are *effective* when its actual outcome is consistent with its intended outcome. Considering that a *policy-schema* “π” declares its intended outcome in its *policy-ends* “π_e”, this component determines a *value assessment framework* “$\mathbb{V}_{\pi}$” by which it is possible to assess the policy with regard of *effectiveness*.

  Let “$\mathcal{S}^{\pi}_{\Omega}$” be the resulting *state of the world* after the enactment of a *policy-schema* “π”. One can say that a policy is *effective* iff the valuation of the resulting state “$\mathbb{V}_{\pi}(\mathcal{S}^{\pi}_{\Omega})$” is *better* than the valuation of the intended state “$\mathbb{V}_{\pi}(\pi_e)$” (Expr. 4.5):

  $$\mathbb{V}_{\pi}(\mathcal{S}^{\pi}_{\Omega}) \geqslant \mathbb{V}_{\pi}(\pi_e) + \gamma_{\text{tolerance, effectiveness}} \tag{4.5}$$

- **Desirability.** Policies are *desirable* with regard to the values of a stakeholder (policy-maker or policy-targets). These values incorporate characteristic expectations (social or individual) to be satisfied in the state of the world. Assuming that a stakeholder “$\mathcal{A}_q$” has its own *value assessment framework* “$\mathbb{V}_{\mathcal{A}_q}$”, it is possible to assess a policy with respect of *desirability* by evaluating its outcomes. In particular, the *policy-maker* that uses the *simulator system* “SIM” will assess the policy effects according to its own *value assessment framework* “$\mathbb{V}_{\text{PM}}$”.

  Let “$\mathcal{S}^{\pi}_{\Omega}$” be the resulting *state of the world* after the enactment of a *policy-schema* “π” and “$\mathcal{S}^{*}_{\Omega}$” a *state of the world* deemed desirable. One can say that a policy is *desirable* for a stakeholder iff it leads to a state of the world whose valuation is sufficiently similar to the valuation of that one considered desirable (Expr. 4.6):

  $$\mathbb{V}_{\mathcal{A}}(\mathcal{S}^{\pi}_{\Omega}) \geqslant \mathbb{V}_{\mathcal{A}}(\mathcal{S}^{*}_{\Omega}) + \gamma_{\text{tolerance, desirability}} \tag{4.6}$$

- **Acceptability.** Policies are *acceptable* when stakeholders agree that the policy outcome is sufficiently similar to the outcome of the most known *favourable* alternative intervention (including non-acting) and, therefore, they are not likely to act to produce any policy shift. Assuming that each stakeholder “$\mathcal{A}_j$” has its own *value assessment framework* “$\mathbb{V}_{\mathcal{A}_j}$”, an additional framework can be defined to assess the policy with regard to *acceptability* (“$\mathbb{V}_{\text{h}}$”), taking the involved stakeholders’ judgements into account “$\{\mathbb{V}_{\mathcal{A}_1}, \ldots, \mathbb{V}_{\mathcal{A}_m}\}$”.

  Let “$\mathcal{S}^{\pi}_{\Omega}$” be the resulting *state of the world* after the enactment of a *policy-schema* “π”. One can say that a policy is *acceptable* iff it leads to a state of the world that stakeholders will not try to change (Expr. 4.7):

  $$\mathbb{V}_{h}(\{\mathbb{V}_{\mathcal{A}_1}(\mathcal{S}^{\pi}_{\Omega}), \ldots, \mathbb{V}_{\mathcal{A}_m}(\mathcal{S}^{\pi}_{\Omega})\}) \geqslant \gamma_{\text{tolerance, acceptability}} \tag{4.7}$$

4.5 Policy-schema

A *policy-schema* is the core of the intervention to improve the state of affairs. It establishes the *policy ends* (i.e. what is to be improved) and the *policy means* (i.e. how is that improvement going to be produced). In more practical terms, a *policy-schema* is composed of two main constructs (Fig. 4.7):

(i) *Policy ends*, that define those desirable states of the world, and are expressed through *indicators*.
(ii) *Policy means*, that aim at driving the system towards deemed-desirable states of the world, and are implemented with *instruments* that affect the social activity.

Notice that, as the *policy domain* is framed within a particular time and spatial context, the *policy-schema* has to fit in these scales (i.e. the time the desirable state is due at and the extension that is concerned with the improvement).

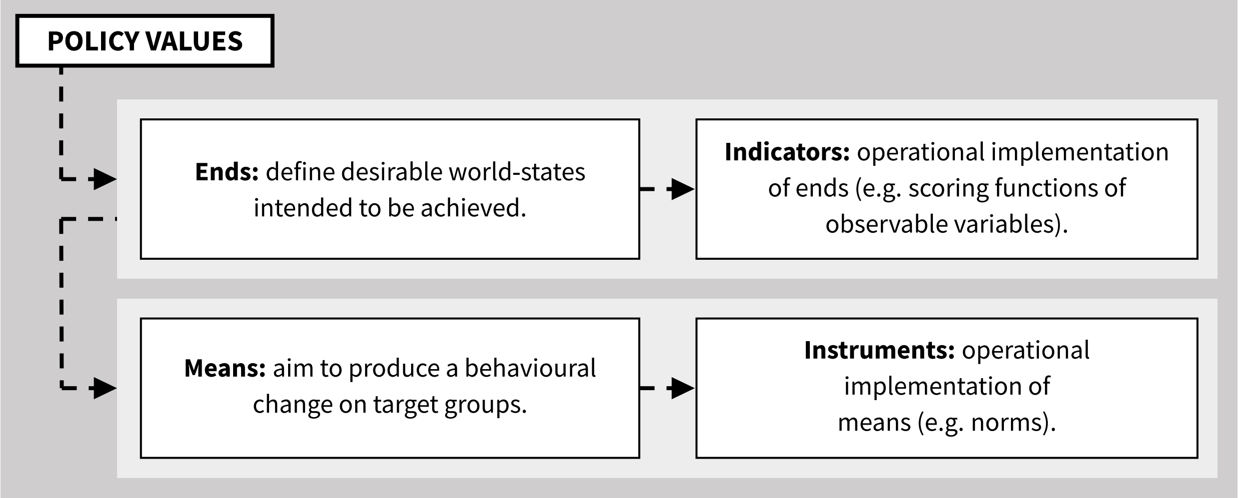


Figure 4.7: Composition of a *policy-schema*. The *policy ends* define the desirable state of the world intended to be achieved with the intervention, relying on *indicators* to do so. The *policy means* aim at driving the system towards that desirable state, by implementing *instruments* in the social space to alter the social activity.

Working definition 13 *A* ***policy-schema*** *“π” is a combination of (i)* ***policy means*** *“π_m”, that are the ways to improve the state of the world, made effective through the enactment of a set of* ***instruments****; and (ii)* ***policy ends*** *“π_e”, that are the definition of those desired states of the world, described by a set of* ***indicators*** *that allow to measure the performance of the policy (Expr. 4.8).*

$$\pi = \langle \pi_m \, , \pi_e \rangle \tag{4.8}$$

As mentioned before, *indicators* are variables (or combination of variables) of the *state of the world* that are said to reflect a particular value (see Wdef. 10). *Instruments* are changes in the institutional framework to change the social coordination system in order to produce a desirable social outcome. Intuitively, multiple instruments may be available for this: financial instruments (e.g. grants, subsidies, loans, insurances, etc.); economic instruments (e.g. prices, taxes, tributes, etc.); regulatory instruments (e.g. laws that afford or forbid certain actions, guidelines, etc.); informational and persuading instruments (e.g. messages, campaigns, advertisement, etc.); social instruments (e.g. networking, distribution channels, social platforms, etc.); and resource/service supply instruments (e.g. infrastructure, technologies, etc.), among others.

4.6 Policy simulator systems to support policy-making

The *simulation environment* is assumed to include a *policy-maker* who is informed by the use of the policy simulator system in order to support policy decisions in the real-world. Using the same terms as in the beginning of this chapter, the policy-maker will design a public policy “Π” to improve a *fragment of the real world* “$\mathcal{W}_{\mathcal{D}}$” and, to do so, relies on the *abstraction* of this fragment of the real world “$\mathcal{M}$” — which includes the *policy-schema* “π” (i.e. representation of the policy)— and its simulation (Table 4.1).

Table 4.1: Abstractions in policy simulator systems to support policy-making

	Reality	**Formal abstraction**	**Simulation**
World	$\mathcal{W}_{\mathcal{D}}$	$\mathcal{M}$	$\Sigma(\mathcal{T}(\widehat{\mathcal{M}})) \rightarrow \widetilde{\mathcal{W}}_{\mathcal{D}}$
Policy	Π	π	

Interestingly, values are reflected by testing an instrument and can be used implicitly by using general heuristics when assessing their raw outcome (see Chapter 9 as example). In this case, it is clear that some resulting *states of the world* are preferred over some others, although a *value assessment framework* has not been explicitly formalised and implemented to process the outcome of the simulation.

Alternatively, policy simulator systems may include an additional layer to deliver a processed simulation output. This means that the simulation output is evaluated by using an implementation of the *value assessment framework* in the simulation platform, thus delivering the overall “social value” of the resulting states (for instance, the framework of the policy-maker, which may include the *meta-values* from Sec. 4.4.2). This can be used to further define “policy problems”, which, using the notion of *design problems* from Simon [236], would be constituted by *means* (i.e. *command variables*, thus *instruments*), *ends* (i.e. *utility function* and *constrains*, thus the implementation of the *value assessment framework*), and *laws* (i.e. *parameters* and *dynamics*, thus the implementation of the *policy model* and *stakeholders*).

4.7 Paradigms when designing policy simulator systems

A *paradigm* is a consolidated socio-cultural understanding that frames collective problems in the real world. In practice, a paradigm shapes the *public policy model* by:

(i) Establishing a definition of what constitute problems and pertinent representations of the *policy domain* (e.g. what are relevant facts, events, causal mechanisms, etc.).
(ii) Imposing *values* and their understandings, as well as those legitimate *policy ends* and associated *indicators* when instituting a *policy-schema*.
(iii) Excluding *means* and *instruments* when instituting a *policy-schema*.
(iv) Recognising individuals or groups as legitimate *stakeholders*.

Chapter 5

Modelling policy shifts

There are additional features for policy simulator systems that can be useful for policy assessment. In particular, policy-makers may want to explore the social acceptance of stakeholders knowing that low levels of such can lead to ineffective interventions or policy shifts. This topic has been identified as key for public policies, as social acceptance issues may jeopardise the achievement of those established goals. Even worse, it may happen that those shifts triggered by the enactment of the policy lead to worse scenarios. To include this feature, the framework has to include additional affordances for the agents in the system. For instance, agents must be able to perceive emergent collective phenomena and react by promoting a political response.

Given this purpose, a proposal for modelling policy shifts in agent-based models is presented in this chapter. This proposal is the basis to model the case study presented in Chapter 10).

5.1 Social acceptance and policy shifts

It has been acknowledged that a key aspect of public policy design is to take social acceptance of stakeholders into account, as it may lead to policy shifts that influence the effectiveness of policies (see [269], for instance). As Botterill and Fenna [37] point out, "the legitimacy of the decision-making process rests in large part on citizens' acceptance of its outcomes, even in cases where they disagree".

Thus, an affordance useful for policy simulator systems is the capacity to test social acceptance. Although causality of social acceptance issues is clearly complex, it is convenient to explore the actual nature of public opinion and citizens' political rationality in order to design interventions to promote social values (see [152]).

Paraphrasing Brunson [44], social acceptability results from a judgemental process by which individuals (1) compare the perceived reality with its known alternatives; and (2) decide whether that reality is superior, or sufficiently similar, to the most favourable alternative condition. If the existing condition is not judged to be sufficient, the individual will initiate behaviour that is believed likely to shift the state of affairs toward a more favourable alternative. Shindler *et al.* [233] identified four reasons to consider social acceptability in management: (i) first, the recognition that management decisions are not only based on technical knowledge, but also values; (ii) second, taking public judgements into account reflects a cooperative attitude; (iii) third, the absence of social acceptance makes difficult to

---

A preliminary version of this chapter was presented in the 20th International Workshop on Multi-Agent-Based Simulation (MABS19) (see [190]).

implement any decision; and (iv) fourth, monitoring social acceptance may help to steer management decisions, since public judgements are subject to change.

With all this in mind, it is also important to understand the information that is available for individuals to assess the state of affairs. Noteworthy, in the information age, *attention economics* has emerged as an approach to the management of information that considers human attention to be a scarce resource. Indeed, people focus on their daily life (i.e. job, family, health, household's economy, leisure, etc.) and therefore may have little time to invest in politics and public affairs. Furthermore, citizens may do not have other relevant resources (e.g. technical, etc.) to process, reason, and store political information of multiple policy domains [3].

This affects how citizens act depending on whether their decision concerns themselves (i.e. individual well-being) or the whole society (i.e. political affairs and social welfare) [74, 3] (see Sec. 3.2.1). Accordingly, citizens are less competent when reasoning about how their decisions affect a complex societal system [3]. As a consequence, people do not engage in well-informed, thoughtful political deliberation. Rather, their political behaviour is generally guided by biases and heuristics (see [152]), such as group loyalties and confirmation biases [3].

Furthermore, it has been explored how citizens derive meaningful policy preferences and attitudes using values, as these are involved in political evaluations [83, 94, 227], moral intuition about social affairs [224] and also in group-identity formation [109, 152]. Accordingly, social acceptance has been related to values (see [105]). In this spirit, De Wildt *et al.* [269] studied social acceptance of energy policies focusing on unsatisfied expectations concerning values, indicating that these could result in issues that could hinder the achievement of policy goals.

5.2 Policy influencers

Another point to consider is who perceives, evaluates, and transfers to citizens the information concerning the whole societal system, making them to take political actions. When competence and expertise are limited, emotions and heuristics play a greater role in reasoning processes, and individuals are more exposed to external messages and persuading mechanisms [74].

It is not common that citizens make political evaluations in particular policy domains directly from raw macro-data, that is, at best, open, trustworthy, and accessible (for instance, from national statistics). In general, most citizens do not have enough resources and skills to process and reason about data that concern multiple policy domains at social scale.

Rather, this information is often provided by *trustworthy* stakeholders. Many of these stakeholders have access to data at a social level and resources to analyse it, which makes them capable of perceiving and reacting emergent social phenomena like demographic changes or economic trends. They provide citizens with information through rhetoric discourses that frame issues and *relevant* data to shape their worldviews and promote action (see [143, 144]). Of course, these actors can act maliciously and aim at manipulating citizens by conveying distorted and biased information, but other cases could refer to actors asking for political action to some crisis (e.g. actions against climate change).

These stakeholders —hereinafter, referred to as *policy influencers*— are usually collectives that represent *political factions* (e.g. mass media, NGOs, think tanks, political parties, interest groups, social movements, etc.), but they can also be distinguished individuals (e.g. celebrities, authoritarian leaders, journalists, a social media account, etc.). Presumably, their trustworthiness and relevance arise from holding similar worldviews and values to the citizen that look upon them. However, have in mind that the *policy influencer* has its own political agenda, and therefore provides useful information with sound framing towards it. In this sense, many scholars have pointed out that politics may be driven by group identities, which are about emotional attachments to an identity instead of adherence to a group ideology [3].

Ultimately, both citizens and influencers assess the state of affairs with regard of their values. Since political participation is motivated by discontent with current situations (see [44, 178]), such evaluations may induce them to take political action.

5.3 Modelling policy shifts

The main idea of this chapter is that among stakeholders there may be influential agents that intervene to advocate their interests and values, eventually producing policy shifts. This extends the conceptual framework of Chapter 4. The intention is to explore affordances for policy simulator systems to allow to model social acceptance and policy shifts.

Namely, this framework-extension considers that some agents may be enabled to perceive, assess and react to an emergent macro-phenomenon (see next section); and that agents are afforded to perform new actions. For instance, some agents can formulate initiatives to change the state of affairs or enact initiatives, while others are only capable to express their support to the initiatives. Initiatives are basically a demand to change the current *policy-schema* and an appeal (e.g. information).

5.3.1 Second-order emergent social phenomena

Second-order emergence social phenomena (*EP2*) refers to the idea that agents may recognise an emerging macro-phenomenon and, as a consequence, intentionally support or hinder the phenomena or the emerging process itself [52, 210, 175]. It contrasts with the (bottom-up) *emergence*, that is the exhibition of macro-structures or properties at the system level as a result of the local behaviour of individuals; and with the (top-down) *immergence*, that is the process through which the macro-level emerging structures modify the micro-level behaviour [210].

For instance, Castelfranchi [52] described *EP2* as the *cognitive emergence* of the macro-phenomena in the agent's mind, and afterwards a process of *cognitive immergence* that changes its behaviour so as to adapt to the situation. Diverse examples were provided to illustrate how the awareness of a phenomenon can promote or discourage it: for instance, in a process of urban segregation, some agents may want to stay close to other agents with similar cultural background, and realise that the territory is being increasingly shared with agents with different cultural background, causing them to adapt their goal to actively oppose new-coming residents in the territory.

Accordingly, one can say that *policy influencers* are able to perceive and evaluate the state of the world at the macro-level, while most citizens can only perceive the state of the world at a local level. Since *policy influencers* may act to support or confront the occurring phenomena, they can advocate for political interventions and gather enough social support to legitimise their proposals. With this in mind, *policy influencers* and *citizens* interact with each other, especially by sharing information.

5.3.2 Implementation of policy shifts

Policy-influencers are stakeholders in the domain —with their own values and goals— that perceive and evaluate the state of affairs at the macro-level. On the contrary, *policy-targets* are not necessarily able to perceive emergent phenomena, but are capable of evaluating the state of the world at the macro-level since they have ethical and political expectations.

That said, *policy-targets* may have no access to raw data or skills to analyse it, but they may receive these from *policy-influencers*. Therefore, there is an exchange of information between *policy-influencers* and *policy-targets*. *Policy-influencers* have their own political-satisfaction models (that includes their social expectations concerning values), which are transferred to *policy-targets* by providing information with a framing discourse. Furthermore, *influencers'* (framed) information is presumably more acceptable if they share the values and the interests of the *policy-target* (in other words, the *policy-target* is biased to consider the *policy-influencer* to be more trustworthy).

With this in mind, *social support* and *policy shifts* may be implemented as follows (Fig. 5.1):

1. The *policy-influencer* agent "$\mathcal{A}_{\text{PI},1}$" has a political-satisfaction model "$m_{\text{PI},1}$" in line with its value profile.
2. The *policy-target* agent "$\mathcal{A}_{\text{PT},1}$" "delegates" its model of political satisfaction to the *policy-influencer* "$\mathcal{A}_{\text{PI},1}$". It receives "$m_{\text{PI},1}$" and adapts its own model "$m_{\text{PT},1}$" to take into account micro-level and macro-level evaluations. In this view, the *policy-target* evaluates its own well-being and also the social welfare.
3. The *policy-target* agent "$\mathcal{A}_{\text{PT},3}$" looks upon two *policy-influencers*, so it receives models "$m_{\text{PI},1}$" and "$m_{\text{PI},1}$". It can take both for its own model "$m_{\text{PT},3}$" (e.g. by combining them, discarding one, etc.).
4. If a *policy-influencer* is not politically satisfied (that is, the *desired world-state* and the *current world-state* are discrepant enough), it may suggest *political demands* "D_k".
5. *Policy-targets* may support these, depending on their own satisfaction and values (for instance, a *policy-target* that weighs its own well-being more than the political state of affairs may not support any *political demand* if it is satisfied with regard of its individual situation).
6. The *policy-maker* "$\mathcal{A}_{\text{PM}}$" designs a *policy-schema* "π" (i.e. *means* "π_m" and *ends* "π_e") according to those deemed-relevant social values and, presumably, taking into consideration the political demands raised in the social space.
7. It is possible that particular demands raised by *policy-influencers* may intervene directly on the social space, bypassing *policy-makers* (e.g. persuading messages to encourage *policy-targets* to adopt social norms). This can be interpreted as new, enacted *means*, for the sake of the simulation.
8. An updated *policy-schema* is eventually enacted in the social space.

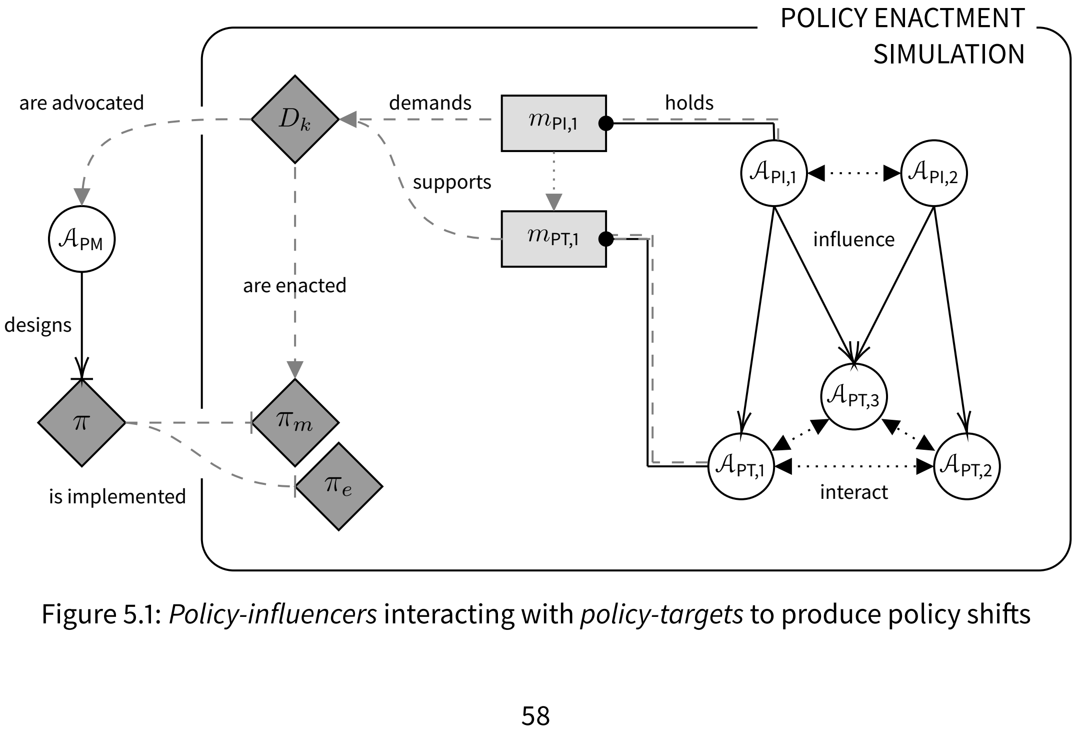


Figure 5.1: *Policy-influencers* interacting with *policy-targets* to produce policy shifts

5.3.3 Observations of this model of policy shifts

Irrationality. Political-satisfaction models are *irrational* when they are *unfeasible*. For instance, an agent that wants to satisfy two values at the same time that are directly opposed, being that situation impossible to be reached. Then, the agent is being irrational because their demands are impossible to be satisfied. For example: *"I want that nobody pays taxes, but I want public education for everybody"*.

It is true, however, that *unfeasibility* is hard to be demonstrated. Indeed, some political actors may invoke unfeasibility as an argument to rhetorically attack other stakeholders (in which case, it would not be *politically unfeasible*, but rather *politically undesirable* from the argument-maker's point of view).

When a *policy-influencer* holds a political-satisfaction model that is irrational, this could lead to a perpetual state of dissatisfaction, no matter the policy enacted. If these models are transferred to *policy-targets*, they are likely to be perpetually displeased too. This would entail unstable social scenarios, as the *policy-influencer* (or another one) could take advantage of the situation to make political demands that do not address the actual problem and consequently do not change the situation, either intentionally (to promote its own agenda) or unwittingly (due to the lack of diverse resources).

Policy local/social contexts misalignment. *Policy-targets* might be incapable of perceiving the attainment of policy ends at the macro-scale, either because they do not receive the information by trusted *policy-influencers*, or because they do not have the values to consider these policy outcomes relevant. Nonetheless, *policy-targets* are aware of the local effects of the policy. If the local effects are viewed as negative (e.g. restrictions or taxes), but *policy-targets* are unable to perceive and value the effects at the macro level (e.g. air pollution reduction), this can lead to social acceptance issues.

Limited competence of *policy-makers*. *Policy-influencers* may evaluate the world-state using indicators that the *policy-makers* in charge may not consider relevant. Therefore, despite sharing the same world-state, they perceive it differently. Consequently, *policy-makers* will receive the political demands as a reaction of the policy effects without being aware of the actual causes of political dissatisfaction. In this case, the governance of the social space may become socially unstable, since it has limited knowledge on what is producing political shocks.

Chapter 6

Ethical concerns when using policy simulator systems

The design of policy simulator systems should take into consideration the ethical concerns that surround them, so that the use of these tools is appropriate. Although agent-based simulation has been acknowledged as a useful methodology to support policy-making [91], it instils a sense of misgiving, as does the use of any type of model for informing policy decisions [19]. Clearly, agent-based models may be based, unconsciously, on errors and artifices, if modelling processes have not been appropriate [87]. Furthermore, models can become “black box” tools [244], informing decisions despite ignoring the underlying assumptions, or they can contribute to crystallise the way to address social phenomena, inhibiting the exploration of alternative models and explanations. In the worst case, such tools may lead decision-makers to the abdication of their responsibility.

Hence, decision-makers and modellers should be aware of the limitations of these tools, if these are to be used to support policy-making. Users and other stakeholders should be informed about the limitations and concerns of simulation for policy-making. Also, designers and modellers should be requested to consider issues that are specific to the modelling of public policies, and not only regard user requirements.

This chapter explores the ethical concerns when using a simulator for policy-making from a value perspective. Although the focus is especially placed on building the policy model at the core of the simulator system (as shown in Chapter 8), other ethical issues are mentioned so as to understand the ethical implications of these systems for policy-making.

6.1 Ethical concerns at three levels

Conceptually speaking, policy-making is an ethical space: policy-makers intend to improve a fragment of the world and design instruments to achieve such improvement. When policy-makers decide to support their decisions with agent-based simulation, they should be aware of ethical concerns in three levels. First, when defining the *use of the simulator*, where policy-makers must decide how they are going to use it to back their policy decisions (e.g. exploration of alternative policies, support arguments to negotiate with other stakeholders, etc.). Second, when *designing the simulator*, which is

---

A preliminary version of this chapter was presented in the 15th Annual Conference of the European Social Simulation Association (SSC19) (together with Chapter 8) (see [191]).

an artifact that will be used to simulate policies and mediate the interaction with other stakeholders. And third, when *building the public policy model*, where policy-makers and modellers create an agent-based model to represent a given policy domain and address a particular problem.

The class of ethical considerations involved in these levels are different but they all essentially amount to value judgements (that is, decide whether something is right or good). Hence, identifying the values involved allow to shed light on how to deal with the ethical concerns when using simulation for policy-making. In this spirit, three levels of ethical concerns are distinguished (Fig. 6.1):[1]

(i) **Top-level:** Use of the simulator system
(ii) **Middle-level:** Design of the simulator system
(iii) **Bottom-level:** Design of the public policy model

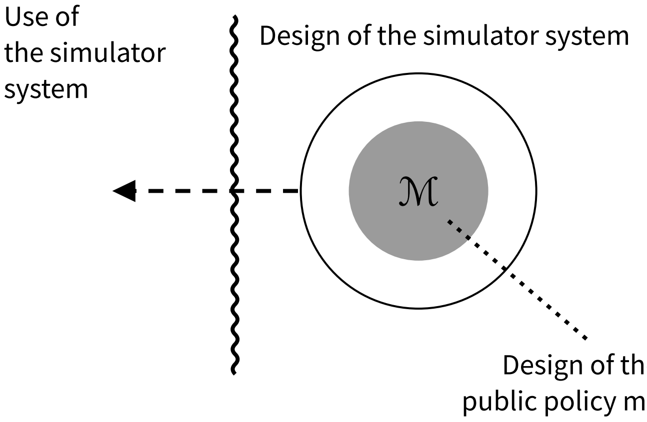


Figure 6.1: Diagram of the three levels of ethical concerns in policy simulator systems

6.1.1 Use of the simulator system

Ethical concerns at this level has to do with the purpose the user is going to use the simulator for. The policy-making cycle involve multiple stages where simulators can be used. To name a few, simulation can be used to explore the effects of different policy instruments, to provide evidence for argumentation and negotiation, and to support policy enactment (that is, to actually *decide* to enact a policy because of firmly relying on the simulation output).

Hence, ethical concerns are related mostly to the conduct of the user with regard of the system. In other words, that the policy-maker is using the system *appropriately*, considering its limitations. Policy-makers must use the simulator bound to the ethical responsibility of their role as public servant. The values involved may be subsumed in the notion of responsibility that should imbue the ethical choices that policy-makers face while using the model during the policy cycle. For instance, the user may consider additional values to assess the simulated outcome despite not having a direct representation in the public policy model (e.g. transparency, responsiveness, accountability, compliance, etc.).

Accordingly, since users do have responsibility, policy-makers must be aware of:

(a) the designers' intended use for the system;
(b) the assumptions and limitations of the model;

---

[1] Notice that these ethical concerns are present in the *simulation environment* (i.e. those elements that anchor the simulation process into the real-world), which includes, among other elements, the *simulator* (with the public policy model), the *user*, and the *experimental setting* (see Sec. 4.1.3).

(c) the possible effects of using the system;
(d) the entitlement of the user with regard to the use that is giving to the system;
(e) the type of knowledge that can be generated with the simulator, and its underlying assumptions;
(f) the change in how policies are designed and implemented that involves the use of computational models, in terms of accountability, reliance and democratic control (see [97]);
(g) the epistemic implications of simulation: the simulator is neither an expert by itself nor it is self-validating. Although simulation may be a reliable source of knowledge, interpretation of its outputs requires to be familiar with theoretical principles and engineering best practices (see [251]).

Design of experiments

Theoretically, simulation allows to explore the system with regard to standards that are difficult to be explored in the real world [23]. For instance, the enactment of a policy in the real world is a one-shot event, since the context will never be exactly the same. In contrast, the policy enactment can be simulated multiple times, maintaining or changing in a controlled manner the conditions of the system. Besides, simulation allows to test the long-term effects of the policy, by comprising the time period or accelerating the passage of time, and do those tests without committing the same amount of resources. Not to mention the ethical implications that would have to test policies with real people, which simulation can easily avoid. Thus, simulation enables an experimental approach in policy science that is difficult to have in the real world (for instance, simulation outputs can be statistically analysed and more easily reproduced).

The simulator system is, in actual terms, a laboratory where to conduct experiments. However, the simulator is neither an expert by itself nor its outputs are self-validating. For this reason, it is convenient that field experts are part of the simulation environment. Experts are familiar with theoretical principles and engineering best practices that underlie the simulator system, and they can appropriately interpret its outputs (see [251]).

The simulator is used by considering an experimental setting, which define the set of experiments, protocols, and techniques to generate knowledge about the real world from analysing one or multiple simulated realities. By tuning such setting, it is possible to assess policies and social systems with regard to additional values (escaping the “one-shot constrains”): robustness, sensitivity, adaptability, confidence, resilience, endurance, redundancy, scalability, etc.

6.1.2 Design of the simulator system

Ethical concerns at this level have to do with how the simulator is going to be designed, taking in mind that it should be a useful artifact for a particular intended purpose. It has to do with ensuring that the system supports, besides its intended purpose, an ethically responsible use for policy-making. Considering the analogy between the simulator and a laboratory, one would want that a chemical laboratory is properly designed and maintained: it is safe, clean, and neat. But also that it has protocols that ensure their appropriate use (e.g. safety rules, emergency protocols, etc.).

Hence, the simulator, as a technical artifact, should be *conscientiously* designed by taking multiple values into account [176]:

(a) Those values related to the social impact of using the system (e.g. liability, ownership, accessibility, accountability, responsibility, etc.).
(b) Those personal values of the users and concerned stakeholders (e.g. recognition, privacy, reciprocity, transparency, inclusiveness, attention, etc.).
(c) Those values related to the concerning the system as a technical artifact (e.g. robustness, correctness, reliability, efficiency, validity, accuracy, etc.).

6.1.3 Design of the policy model

Ethical concerns at this level have to do with how the policy model is abstracted. As mentioned in Sec. 3.1.2, modelling processes imply choices that have ethical implications with respect to the domain of interest. In plain words, these decisions define how the policy domain is represented, by committing to a specific domain ontology that determines the variables, the set of events and actions that may modify the state of the world, and the characterisation of the agents that populate the model (especially, the definition of their (value-sensitive) decision-making models).

Hence, a key ethical issue is to choose the relevant values in the policy domain and to make them operational by identifying those variables that are pertinent for each of those values. Then, some of these variables are chosen as indicators of values to assess social outcomes. This has implications in policy-making, because some actions and events are identified to change such indicators, since they are essential when modelling the pertinent interventions on the social system.

In general, values are used to evaluate the (resulting) states of the world and to evaluate whether some interventions lead to a *good* state. With this in mind, there is the need to make these ethical aspects operational. For this reason, a *consequentalist view of values* as explained in Sec. 3.2.2 is particularly useful. It says that a value is defined by means of indicators that can be observed in the world at any time. It makes possible to say that one state of the world is better than another —with respect to a value— provided that it has a better scoring for the indicator of that value, and to decide whether an action is better than another —with respect to a value— by comparing their effects on the state of the world.

That said, it can be approached as a “representation problem”, aiming at translating abstract universal values that should be considered in the domain into observable factual indicators that reflect these values appropriately and therefore allow to design interventions adequately. This process is illustrated for the water domain in Chapter 8.

Part III

Study domain: Water

Chapter 7

Introduction to the water domain

Water is the most essential resource on Earth. It is indispensable for life and it is involved in many facets of human civilisation (i.e. agriculture, public health, energy generation, culture and religion, etc.). Human communities have withdrawn water from the nature to meet their needs and then returned it to the environment for millennia —with a lower quality. Not only is water used for basic supply, but also it is indispensable for sustaining ecosystems that provide wide range of services that are essential for human well-being [164]. Consequently, a proper water governance is fundamental to guarantee the sustainability and the availability of water resources and services in the future.

Since water management has an impact across the interwoven socio-economic and ecological systems, its governance is an issue in which many groups and communities have different interests. Not surprisingly, water is at the heart of several social and ecological conflicts along the history of humankind.

7.1 Water as an ethical space and policy domain

Water governance implies value-laden decisions and ethical judgements of many sorts, as they are necessary to face the multiple trade-offs that arise in water collective problems [181, 99, 54]. Therefore, water governance establishes a state of affairs deemed as desirable, and regulates the management of water according to it by means of a system of political, social, economic, and administrative components [181]. Although its vital importance and political implications, water policy is not a popular topic itself, often subordinated to economic matters —whose paradigms are being questioned nowadays [161, 84]. As Meisch [162] pointed out, water governance has to deal with those called "wicked problems", that are complex situations for which there is neither a simple nor a single correct solution; therefore, "social solutions cannot be judged by the standards of true or untrue, but only of better or worse. In this sense, solutions to wicked problems are always of a political nature because they are intimately related questions of how water-human relationships in a certain location should and can be shaped at a certain time—according to affected citizens' notions of the good life and the right actions in relation to their water. Due to dynamics within hydro-social cycles or "waterscapes", these relationships will never be finally settled, but must always be socially renegotiated and ethically justified".

These issues are particularly urgent nowadays, since water is at the core of the sustainability challenges that human civilisation is facing [204] and of multiple Sustainable Development Goals —which are to contribute to the "peace and prosperity for people and the planet, now and into the future". Global conditions are changing rapidly and they pose a severe threat to socio-ecological systems [48, 204], especially in the water domain, as they are expected to compromise water security and hence social welfare [232, 254]. Pressures are of various kinds (e.g. urbanisation and population growth, ageing infrastructure, global warming, technological disruption, etc.) and they are challenging water governance from many fronts [232, 165]. Thus, such threats are increasingly severe for urban areas, as they concentrate 55 % of the global population today, rising up to 68 % by 2050 [259], and where water crisis are already occurring (see [51, 266, 241]), besides being other social conflicts such as gentrification, unemployment, and citizen insecurity more pronounced. In addition, cities are where socio-technological innovations land first, usually without knowing their long-term consequences. What is undoubtedly true is that water conflicts in the era of climate change —together with technological disruption and political upheaval— are going to be aggravated due to a fiercer competition and social unrest.

Water crisis are approached as crisis of governance in many institutions [258, 54]. Accordingly, it is becoming critical to reshape water governance in order to adapt the water use and management to a new era for ensuring social well-being. With this in mind, it is convenient to consider the socio-cultural context of water systems, and explore new approaches to support policy exploration and assessment, especially those issues that concern social coordination. Consequently, addressing collective water problems effectively depends on enhancing the understanding of individual/collective action to support decision-making in those matters. In this spirit, a new field called socio-hydrology [239] aims at studying the water cycle together with a socio-cultural perspective [66], for which AI-based approaches (and, in particular, agent-based modelling) may be extremely meaningful.

Adopting a value-based perspective may illustrate how water governance and water management impact on aspects that are subject to be *valued* from a social perspective. This can potentially lead to better decision-frameworks and analysis-tools in the water sector. Although water management has traditionally focused on *efficiency* (i.e. eventually translated into economic costs per unit of water), it must be understood that many other social values are present in it, which can lead to social conflicts if not recognised properly. It is therefore necessary to determine what values are relevant and how they can be expressed, in order to be properly incorporated in strategical decisions and in socio-technical developments.

7.2 Values in the water domain

7.2.1 Socio-hydrology

Water governance and management are connected to many universal human needs and public values. Water is related to public health, to food and energy production, and to social justice and progress. Not surprisingly, water governance is often viewed as a moral duty of the modern social democracies —which is why water has been traditionally managed by means of public services in Europe.

Water policy decisions are expected to be more effective if they are based on an accurate understanding of the human values and social dynamics involved (see [239, 243]). For water governance to face sustainability challenges it is required to understand how societies and water systems interact and what human communities value [209]. In this spirit, the field of socio-hydrology suggests that the water cycle cannot be understood without considering human culture, values, and institutions (and vice versa) [239, 146].

Since governance and policies do reflect values of stakeholders [33, 99], water governance has to do with political processes in which actors advocate their own values and worldviews, thus determining how water issues are framed and addressed [54]. Accordingly, water governance decisions involve ethical judgements on which states are desirable, how the trade-offs should be solved, and how agents should be guided [43].

7.2.2 Values

Values have been considered widely in water, but they have been addressed quite differently. The research of values in water corresponds greatly with the research of environmental values (see [67]).

In these works, values are understood as legitimate uses or interests *assigned* to water (e.g. agriculture, energy generation), as well as other services of water ecosystems (e.g. climate regulation) [157]. Given this view, the *assigned value* of water refers to the contribution to meet a specific goal or objective [60]. Values are also covered in water ethics, concerning the inherent universal values that water has (e.g. spiritual, aesthetic, etc.), usually by adopting an objective moral perspective [99].

Held values are drivers of the decision-making and goal-setting of stakeholders, eventually shaping the social outcome and impacting on human well-being and ecosystems' health [209]. This category is aligned with the notion of values presented in the Schwartz Theory of Basic Values [219]. Typically, the *held values* are grouped by (i) self-interest or egoistic values (e.g. personal profit); (ii) humanist-altruistic values (e.g. community well-being); and (iii) biosphere-altruistic values (e.g. ecosystem health).

Values are also covered in water ethics, concerning the inherent universal values that water has (e.g. spiritual, aesthetic, etc.) that should guide water governance [99]. In these works, values of water are recognised by adopting an objective moral perspective.

Other approaches understand values as principles that should be endorsed by water governance and that are involved in defining ends and hierarchies [218, 100], like *transparency* or *social justice*. There are diverse frameworks that have been proposed to evaluate water governance concerning multiple values (e.g. [268, 262, 8]), but typical values concern water *availability*, *affordability*, *accessibility* and *safety*. This view reminds the notion of *public values* as consensual expressions of what is desirable in society and governance (see [39]), but focusing particularly on the water domain.

The aim of this dissertation is neither to define how to understand values in water nor state what values water governance should promote and preserve. Rather, the point is to explore how values can be imbued in policies and considered in computational models of public policy. Noteworthy, values were already used to guide policy design for water utilities [134]. Nonetheless, this dissertation aims at expanding this topic by considering how this can be approached in computational decision-support systems based on simulation.

7.3 Water and agent-based simulation

A relevant line of research for water policy-making takes the socio-hydrology approach [239]. This field focuses on modelling the coupled human-water system, explicitly considering social and water dynamics together [239]. Although equation-based models are more popular in socio-hydrology, agent-based simulation is becoming a relevant modelling approach in the field [140]. Indeed, agent-based models are suggested to overcome many limitations of equation-based models, as heterogeneity of agents or explicit social interaction of agents [261]. In this view, agent-based modelling can combine human behavioural models with biophysical models to focus on water issues.

Accordingly, agent-based simulation has been widely used to explore human social behaviour water-related domains. For instance, it has been used in the agricultural domain [118, 31, 217, 27, 117], focusing typically on land-use changes [182, 113] and on the implementation of payment for ecosystem services [231, 56]. It has been also applied in water resources management [32] and urban water demand [89, 141]. In environmental studies (not only focused on water), agent-based models have been broadly used to study socio-ecological systems [96, 17].

Chapter 8

Bringing values into computational models: application in the water domain

In Chapter 3, a value-driven approach for modelling was argued to potentially benefit policy simulator systems, because it can make modellers and policy-makers aware of the spectrum of effects that policies may have in multiple aspects.

This chapter aims at providing basic guidelines for introducing values into (agent-based) policy simulator systems. In particular, the approach is illustrated for the water domain. However, it is presumed that the methodology is general enough to be useful for bringing values into models situated in other policy domains.

8.1 Values in computational models situated in policy domains

Values are abstract constructs that transcend specific situations [219]. Since simulator systems are situated in a particular policy domain, values also require to be situated in that domain and take a specific meaning in the (computational) policy model. That said, there are two relevant questions for the policy-modelling process: *"what are the relevant values in a particular policy domain and a specific policy problem?"* and *"how to make these values operational in the computational model?"*.

The conundrum is to translate abstract values into an entity that satisfies some requirements: first, policy interventions and actions must be able to change their *state* (otherwise, policy-making cannot *improve* the *state of the world*); second, they must be *observable* in the social space, for agents to be capable to *assess* their *state*; and third, they must be able to be *coded*, for computational models to be able to simulate the policy effects on their *status*. In sum, value-driven policy simulator systems must enable to observe, measure and monitor *values*.

Assuming a consequentialist view of values, as explained in Sec. 3.2.2, is specially helpful at this point, since abstract values can be translated into *factual indicators* that stand for them, satisfying the aforementioned requirements. *Factual indicators* are nothing but variables of the *domain ontology*, whose social meaning reflects the values of interest. Hence, the value-instantiation process goes through different stages (Fig. 8.1): (i) from abstract values, (ii) to their contextualisation in the domain, (iii) to factual indicators in the computational model (thereby anchoring abstract values to the objective simulated world). With this in mind, *values* must be conveniently present during the modelling process,

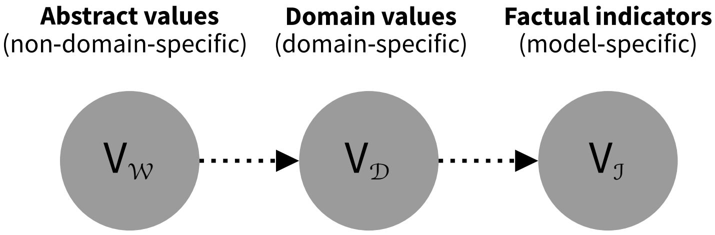


Figure 8.1: Value instantiation process for computational models

as they may restate the *domain ontology*.

Notice that this process may be quite recursive, and it may be not clearly structured. Nonetheless, it is basic for building a policy model (despite not being explicitly conducted).

8.1.1 Abstract values relevant in the policy domain

Values are present when creating the policy model that will support decision-making. Above all, values determine what is *relevant* in the domain for a collective problem and underpin the assessment of policies. Thus, choosing those relevant values in the policy domain is the first fundamental issue when building policy simulator systems.

One conclusion is drawn from the works and theories on values reviewed in Chapter 2: the compilation and practical definition of values that drive decision-making in public affairs is quite difficult. In other words, stakeholders and public servants may have different understandings of what constitute *social values* in a particular domain and what they mean. Thus, different stakeholders may invoke the same *value* to justify decisions whose consequences are radically disparate. One reason for this is, simply, that stakeholders have different interests and beliefs, and this is translated into different social expectations for the same value.

Selection of relevant values in the water domain

There may be different proposals about which values should be considered in the water domain. For the sake of example, and only for pedagogical use, six social value items are taken from the works reviewed in Chapter 2:

- **Efficiency.** Making use of resources in a responsible way.
- **Competence.** Proving beneficial to those who have invested resources in the service.
- **Sustainability.** Proving beneficial to society in the future.
- **Security.** Ensuring that society is safe from threats from within and without.
- **Modernity.** Using and promoting the newest and best approaches and technologies.
- **Social justice.** Concerning for the welfare of the larger society, specially those who are at a disadvantage.

8.1.2 Contextualising values in the policy domain

The following stage is to situate those abstract values in the domain which allows to speak in more concrete terms. That is to say that they have to be reformulated to fit the water domain and make

---

A preliminary version of this chapter was presented in the 15th Annual Conference of the European Social Simulation Association (SSC19) (together with Chapter 6) (see [191]).

them able to have an operational translation in computational models for simulation. This step is better illustrated with the example:

Contextualising values in the water domain

Since water is present in many activities (see Chapter 7), the contextualisation process approaches the water domain from a general point of view:

- **Efficiency** understood as "*making a responsible use of natural resources*". Under a scenario of global change that is increasing pressures on water systems as well as water demand (e.g. climate change, population growth, etc.), it is essential to use water resources *efficiently*. Although means to promote *efficiency* are typically based on technological solutions, they can also be based on social interventions, such as encouraging citizens to not waste water or to accept reclaimed water.

  Several directives and frameworks aim at promoting *efficiency* in the water domain. For instance, the European Commission has outlined the Roadmap to a Resource Efficient Europe to transform Europe's economy by 2050, aiming at increasing resource productivity while decoupling economic growth from resource use and its environmental impact [77]. In the same vein, the *circular economy* paradigm is being promoted in Europe, strategy that aims at closing material and energy cycles. It is considered that the traditional *linear economy* has proved unsustainable, because ecosystems cannot cope with the demands for resources and the reception of emissions and waste. In contrast, in the circular model energy and material flows are planned as minor cycles that practically close the production system (e.g. repair, recycling, etc.), in addition to consider any waste as resources for other industrial processes (which is also known as *industrial symbiosis*). In the case of the water circular economy, reuse and reclamation of water resources take special importance.

  Efficiency in water domain may also take a broader approach by considering the fact that water is extremely related to other resources, as describes the Food-Energy-Water nexus [46] (e.g. the energy used to desalinise water). Therefore, seeking to "improve" *efficiency* in the water domain may look to intervene on other relevant resources (e.g. fuels, nutrients).

  Furthermore, *efficiency* can focus on distribution of resources. In this view, *allocative efficiency* is understood as allocating water resources to those uses and users that generate the most value —which, as mentioned in previous chapters, depends on how *value* is understood. Since *efficiency* is typically framed with the notion of maximising the economic output, the focus is often on increasing productivity; however, some scholars point that *efficiency* may be promoted in order to decrease the input of resources [193, 150].

- **Competence** understood as "*contributing to the best use of economic resources of water services, in order to guarantee standards of service and costs that are appropriate to the levels of service provided*". Water services address multiple societal needs, and are supported through coercive taxation. Therefore water services should be properly managed to ensure their sustainability to be able to address those societal needs in the future. In this line, the Water Framework Directive (WFD) (Directive 2000/60/EC) established the principle of recovery of the costs of water services including environmental and resource costs. Accordingly, two types of costs are distinguished (i) financial costs, which include operation, maintenance, and investment costs (i.e. capital and interests); and (ii) environmental and resource costs, derived from the effects that water uses have on the health of water ecosystems. The aim is that water-pricing policies provide adequate incentives for users to contribute to the environmental objectives of the Directive, and that these guarantee the recovery of costs of the multiple water services —taking account of a fair distribution of costs (e.g. polluter pays principle, by which the cost of implementing environmental mitigation and restoration measures should be covered by the one who caused such environmental deterioration). Multiple strategies can fit this value: water utilities may focus on appealing to citizens to be "responsible" and to not perform behaviours that will increase costs unnecessarily (e.g. not flushing wet wipes because they block sewers); or AI techniques can be used to mitigate high-cost events (e.g. mitigate peak demands by implementing dynamic pricing).

- **Environmental protection** understood as "*preserving, protecting, and improving the health of ecosystems*". Given that ecosystems provide essential services to human beings (e.g. freshwater provision, flood protection, etc.) [164], it is convenient to improve, preserve, and monitor their quality and equilibrium. It is well-known that many human activities are deteriorating the quality of ecosystems; for instance, the intensive nitrogen fertilisation to increase crop production, combined with the destruction of riparian vegetation (which acts as filtering mechanism), causes that the excess of nutrients can not be retained and recycled locally, and is transported by runoff into aquifers, rivers and lakes, and finally into oceans [240]. The excess of nutrients can cause eutrophication in water ecosystems, a phenomenon that consists in the excessive growth of algae, which leads to oxygen depletion and the death of numerous organisms, with severe consequences for biodiversity. Notably, enhancing the status of water bodies was one of the main objectives of the WFD at the time it was enacted, for which it established that Member States were required to implement River Basin Management Plans (RBMPs).[1] Given that purpose, the WFD established a methodology to assess the status of water bodies, setting indicators for monitoring diverse relevant aspects (i.e. quantitative and chemical aspects for groundwater bodies, and biological, hydromorphological, and physico-chemical aspects for surface water bodies).

- **Water security** understood as "*ensuring that the population is safe from water-related threats*". Water-related threats have to do with loss of water availability, floods, and droughts. Eventually, water (in)security may lead to disease spreading, food insecurity, and political instability [93]. Accordingly, *water security* is significantly related to public health, for which is needed a safe supply of good quality water and appropriate sanitation of wastewater for the population.

  Although water is one of the most essential resources, human activity is severely affecting water security. For instance, as mentioned before, nitrogen pollution may affect aquifers due to an inappropriate agricultural management, making the provision of water resources practically impossible, since groundwater pollution is difficult to reverse. The intake of water with high concentrations of nitrates causes health problems, such as methemoglobinemia in infants, in addition to other chronic diseases [274]. For this reason, the European Directive concerning the protection of waters against pollution caused by nitrates from agricultural sources (Directive 91/676/EEC) requires the designation of vulnerable zones, identifying those water bodies water that may be affected by infiltration and runoff whose nitrate concentration exceeds 50 mg/L.

  Notably, climate change is altering the water cycle and, as a consequence, it is increasing the risk of water-related threats. For instance, a general decrease in water resources is occurring in Spain due to the impacts of climate change [55]. In addition, climate change increases the likelihood and magnitude of extreme events. Thus, extreme precipitation events will be more frequent and greater magnitude, making more difficult to store water resources and also increasing the risks of floods —which is direct threat for people, but also indirect considering that

---

[1] The Commission assessment of the second River Basin Management Plans was reported in February 2019 (SWD(2019) 30 final).

water infrastructures may not be designed for such events (e.g. sewer flooding, damages in supply networks, etc.). Climate change also exacerbates drought risk, negatively affecting drinking-water supplies and food production, among other issues, which may lead to harsh competence and migration. In Spain, the territory that will suffer extreme drought will also be greater, aggravated in summer by the seasonal change in fluvial flow and the reduction of water reserves in the form of snow in the mountains [55]. In fact, severe water-threats are already occurring in cities (e.g. Cape Town [266] and Sidney [241]), where 55 % of the global population concentrates nowadays —rising up to 68 % by 2050 [259].

— **Modernity** understood as "*using and promoting the newest and best approaches and technologies in water use and management*". In this view, modernisation consists of a transformation process, by which a traditional community (often perceived as backward), becomes a modern, recent-time community. Hence, innovation and technology adoption become a value by itself. For example, the Spanish National Irrigation Plan Horizon 2008 focused on rural development by *modernising* much of the extension that was used for irrigation agriculture —meaning that traditional systems based on open channels distribution and flood irrigation are replaced by pressurised networks and sprinkler or drip irrigation— since irrigation agriculture has been traditionally seen as a symbol of prosperity [148]. In the same vein, the water policy paradigm in Spain during the military dictatorship (1939–1975) focused on the construction of reservoirs and dams to increase the supply of water resources, thereby symbolising the development of the country [149]. In more recent times, ICTs (e.g. digital water-meters, Internet of Things, 5G), AI techniques (e.g. predictive analytics using *machine learning*), and related marketing-based concepts (e.g. *digital twins*, *smart cities*, *industry 4.0*) are being promoted in the water sector to, presumably, improve the management of water (See [97]).

— **Social justice** understood as "*ensuring that the right of citizens to access a safe water supply and sanitation is protected for them to live a decent life, distributing costs by taking the capacities of the diverse people of the society into account*". In this line, the United Nations Member States adopted the 2030 Agenda for Sustainable Development in 2015 in order to contribute to "peace and prosperity for people and the planet, now and into the future".[2] It established 17 Sustainable Development Goals (SDGs), many of them related to aspects of *social justice* (e.g. SDG1: No Poverty, or SDG2: Zero Hunger). There are goals that are specially relevant in the water domain, such as the SDG6: Clean Water and Sanitation, which focuses on the fact that there are people who still lack access to safe water supply and sanitation. Also regulations and laws may address these issues. For instance, in Catalonia (Spain) a law was passed to address the emergency in housing and energy poverty (i.e. lack of access to basic services as energy and water supply) (Law 24/2015, of July 29). This law in particular establishes that public administrations must ensure that water, gas and electricity supply companies provide grants or apply discounts to those vulnerable families. Furthermore, the company must contact with the municipal Social Services to determine whether the family is at risk of social exclusion prior to interrupt the supply to any household.

Noticeably, *contextualised values* are closely entwined. For instance, *environment protection* is often pursued because of *water security*, and *modernity* by technological development is often made to increase *efficiency*. This points out that values, concerning a complex system, are hard to be instantiated as isolated realms. However, although situating these values leads to highly related understandings, they express different motivations when justifying policy decisions, thus expressing different *policy ends* and relying on different *policy means*. Furthermore, the fact that such realms are linked is the

---

[2] `https://sustainabledevelopment.un.org`

reason why there exist value trade-offs in policy decisions. Hence, the gain in one value may be at expense of losing in another other value.

8.1.3 Instantiation of values into factual indicators

Once the contextualisation is made, one needs to define the *factual indicators* that anchor the values to the simulation. The instance of these indicators defines the *state of the world*.

Hence, the choice of *factual indicators* is key for policy-making. First, because they determine the *ontology* of the model, as mentioned before. Second, because they preset the relevant policy instruments as those instruments that have an impact on those indicators. Third, expert knowledge on the domain is necessary (see Sec. 3.1).

One may follow a recursive approach to this selection process: (i) start with a list of contextualised values, propose a list of indicators; (ii) for each indicator, find policy instruments and actions that involve it; (iii) identify missing indicators in these instruments, and update the list of indicators until no new indicators are found; (iv) and then from the indicators back, update the set of instruments, and also consider if necessary to update the list of relevant values with new items that are related to those new elements.

It is true, however, that some directives, roadmaps, and policy frameworks propose particular indicators and instruments. For instance, the Roadmap to a Resource Efficient Europe considers, for the water domain, particular measures (e.g. smart metering, guidelines for water re-use, etc.) and indicators (e.g. leakage in water infrastructure, water savings in irrigation, etc.) [77]. These resources may provide a guiding modelling framework.

Interestingly, any instantiation of values into *factual indicators* is questionable. Nonetheless, this is an inherent issue when working with values: every stakeholder may define and understand values differently in the domain, due to contrasts in culture, experiences, education, interests, and so on. Not surprisingly, values are contested and disputed in political arenas. However, this is not necessarily a downside —in fact, it may be a virtue: since the world is constantly changing (i.e. society, science, technology, politics, etc.), values can be redefined in new contexts and situations. Policy problems are 'wicked problems' [162], which means that there are no unique correct solutions.

Defining *policy-schemas* in the policy domain

Once those relevant values have been translated into *factual indicators*, these can be used to define policy interventions. That is, indicators are used to model *policy-schemas* operationally in agent-based models (see Sec. 3.1.5): the *ends* of a policy, that need to be expressed in terms of *indicators*, and the *means* to achieve them, that are made precise in the form of *instruments* (Fig. 4.7). In this view, the *policy-schema* aims to *improve* the domain-situated social system with respect to some *values* by intervening in those linked indicators. Noteworthy, choosing *ends* has to do with deciding what to prioritise, which is clearly a political task.

If the *policy-schema* has been generated from the model ontology, it should be compatible with the computational model, and therefore able to be *implemented* and *executed* to simulate its effects. For instance, an *instrument* that is based on *"sending messages to make citizens aware of good practices for saving water at home"* requires that the ontology includes citizens, water conservation practices, and some decision-making model focused on adopting such practices. Likewise, projecting the ends onto the indicator "*water use in* m^3 *per month per person*" needs that the ontology includes appropriate variables to compute it (e.g. *volume of water used*, *number of urban dwellers*, etc.). Since the potential instruments must be included in the simulator system, they must be properly modelled, especially their effects, which requires to have scientific knowledge in multiple disciplines in order to do so appropriately.

Instantiating values as indicators and definition of *policy-schemas* in the water domain

Given the *contextualised values* of Sec. 8.1.2, Tables 8.1 and 8.2 show some *factual indicators* (i.e. domain variables and combinations of these) involved in these value items, and how one can identify some generic means to address these value goals and how these are instantiated into instruments that will be brought into the *policy-schema*. Besides, to further illustrate this data structure to model policy proposals, Fig. 8.2 represents a *policy-schema* to improve *water security*.

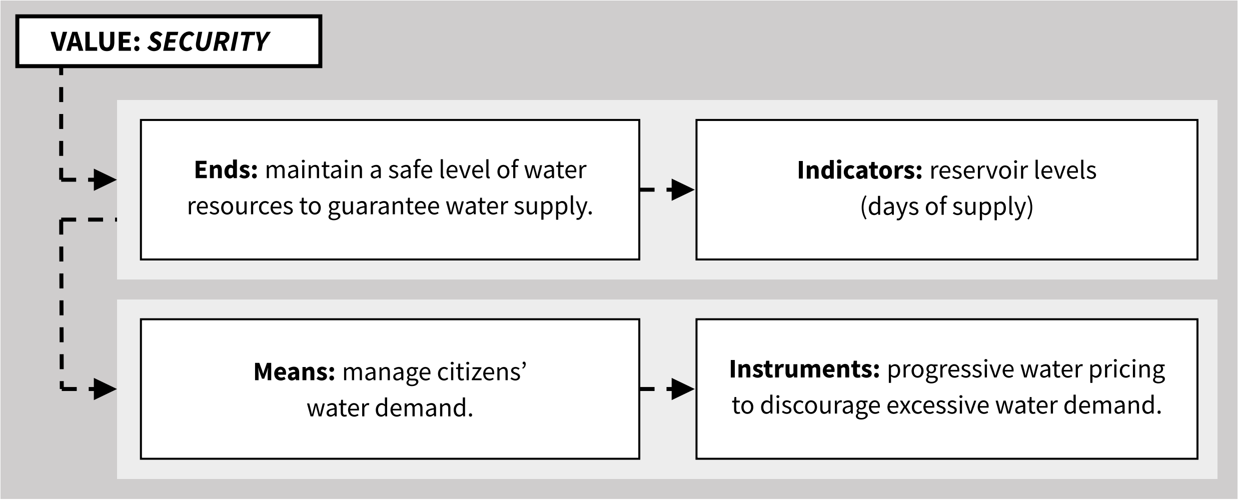


Figure 8.2: Example of instantiation of values into a specific policy-schema. Since *security* was contextualised in the domain as *water security* (thus potentially relating to issues as public health and sanitation, extreme events, food production, political instability, etc.), it is possible to define indicators that reflect the value (which can be measured and computed) and, afterwards, a *policy-schema* aiming at *improving* the state of the world with respect to that value.

Table 8.1: Value instantiations and examples of options to define *policy-schemas* aiming to imbue those values

Value item	Ends	Indicators	Means	Instruments
Efficiency Making a responsible use of natural resources	- Water economy that maximises the economic output - Water public services *responsibly* used by citizens - {...}	- Water use (m^3/month·person) - Share of reused water (%) - Water network losses (%) - Economic efficiency (eur/m^3) - {...}	- Promote circular economy as a new paradigm - Maintenance of existing water infrastructures - Addition of ICTs to infrastructures to improve management - {...}	- Messages to make citizens adopt best practices to save water at home - Plan replacement and maintenance of pipelines - Adoption of smart water-meters in households to detect leakages - {...}
Competence Contributing to the best use of economic resources of water services, in order to guarantee standards of service and costs that are appropriate to the levels of service provided	- Financial sustainability of water services - Water service with a high-quality customer service - {...}	- Infrastructure Operation and Maintenance costs (eur/m^3) - Service-cost recovery rate (%) - Non-revenue water (%) - Water supply interruptions (min/person) - {...}	- Addition of ICTs to water infrastructure to improve management - Control users' water demand - Prevent users from performing neglectful practices - {...}	- AI agent to predict water supply interruptions - AI agent to apply dynamic pricing on water fees in order to avoid peak demands - Messages to appeal citizens to not flush wet wipes - {...}
Environmental protection Preserving, protecting, and improving the health of ecosystems	- Better status of water bodies - Minimisation of environmental impact of stages of urban water cycle - {...}	- Chemical and ecological status of surface water bodies - Discharged-water quality parameters (e.g. BOD, COD, nutrients, etc.) - Contribution to global warming (g CO_2-eq/m^3) - {...}	- Compensate water users for beneficial practices - Best technologies in industrial processes to mitigate environmental impacts - {...}	- Payment for Ecosystem Services (PES) for upstream farmers who protect riparian vegetation in their fields - Subsidies to encourage the use of Best Available Technologies (BAT) - {...}

Table 8.2: Value instantiations and examples of options to define *policy-schemas* aiming to imbue those values (cont.)

Value item	Ends	Indicators	Means	Instruments
Water security Ensuring that the population is safe from water-related threats	- Safe level of water resources - Low detrimental health impacts caused by water sanitation - {...}	- Reservoir levels (%, days of supply) - Mortality- and morbidity-quantifying indicators (e.g. DALY) - Risk of flooding (i.e. probability of occurrence and damage) (eur) - {...}	- Diversify water sources - Control water quality - Control users' water demand - {...}	- Regulation norm that obliges farmers to use reused water - Implementation of environmental standards for water to be discharged - Progressive water pricing to discourage excessive water use - {...}
Modernity Using and promoting the newest and best approaches and technologies in water use and management	- Minimum level of adoption in a community - Strong collaborative network of collaboration between researchers and industries - {...}	- Technology adoption rates (%-time) - Patents (n°/yr) - Academic impact (n°articles/yr, n°cites, etc.) - {...}	- Promote technological adoption - Foster R&D expenditure - {...}	- Subsidies to encourage irrigation technologies - Provide fiscal incentives to get involved in research projects - {...}
Social justice Ensuring that the right of citizens to access a safe water supply and sanitation is protected for them to live a decent life, distributing costs by taking the capacities of the diverse people of the society into account	- A minimum level of redistribution of wealth - Equal opportunities - {...}	- Vulnerable households due to low income (%) - Relative (water) utilities cost on household income (%) - Population without basic access (%) - Households whose water use is under 90 L/p·d (%) - {...}	- Financial social aid for vulnerable households - Water service funding through taxes taking the impact of the user into account - Addition of ICTs to monitor water users - Train citizens to improve their capacities and social inclusion - {...}	- Subsidies for those households whose income is below a minimum level - Specific tariffs for large and industrial users - Smart-meters to detect anomalous low water use - Scholarships for training youth to have jobs in the water sector - {...}

Chapter 9

Case study: Modelling contingent technology adoption in farming irrigation communities

Of all the uses of water, agriculture is the one that requires the greatest proportion of resources worldwide. Consequently, it is a salient subject for environmental policy-making, and adoption of modern irrigation systems is a key means to improve water use efficiency. In this chapter it is presented an agent-based model of the adoption process —known as "modernisation"— of a community constituted by farmer agents. The phenomenon is approached as a contingent innovation adoption: a first stage to reach a collective agreement followed by an individual adoption decision. The model is based on historical data from two Spanish irrigation communities during the period 1975–2010.

9.1 Case study for the conceptual framework

Rural socio-economic development has been a recurrent topic of Spanish public policies which has been traditionally linked to irrigation agriculture, understood as essential part of the agro-industry. Hence, the national hydraulic policy of the past twentieth century was based on increasing the supply of water resources and the transformation of irrigation, in order to generate employment and curb rural depopulation [148]. In this view, modernising irrigation is aimed at achieving territorial balance by competing with the urban regions, which usually attract human capital from the rural world.

This case study is approached as a "simple case" for the sake of illustrating the conceptual framework. This chapter can be seen as it presents a policy simulator system that is used by policy-makers and experts for testing one particular policy-schema that aims at imbuing one particular social value to the social system: *modernity*. In this connection, *modernisation* is understood as a transformation process by which a traditional community (often perceived as backward) becomes a modern, recent-time community. As a consequence, the diffusion and adoption of high-tech irrigation systems is seen as desirable (i.e. *policy end*) and therefore should be promoted, namely by means of economic incentives (i.e *policy means*) (Fig. 9.1). Table 9.1 provides a summary of the case study.

This case study is considered a simple case because there are no value trade-offs nor competing stakeholders that hold different political interests (potential value conflicts are mentioned in the discussion

---

This case study was published as article in the *Journal of Artificial Societies and Social Simulation* (see [189]) (Open access).

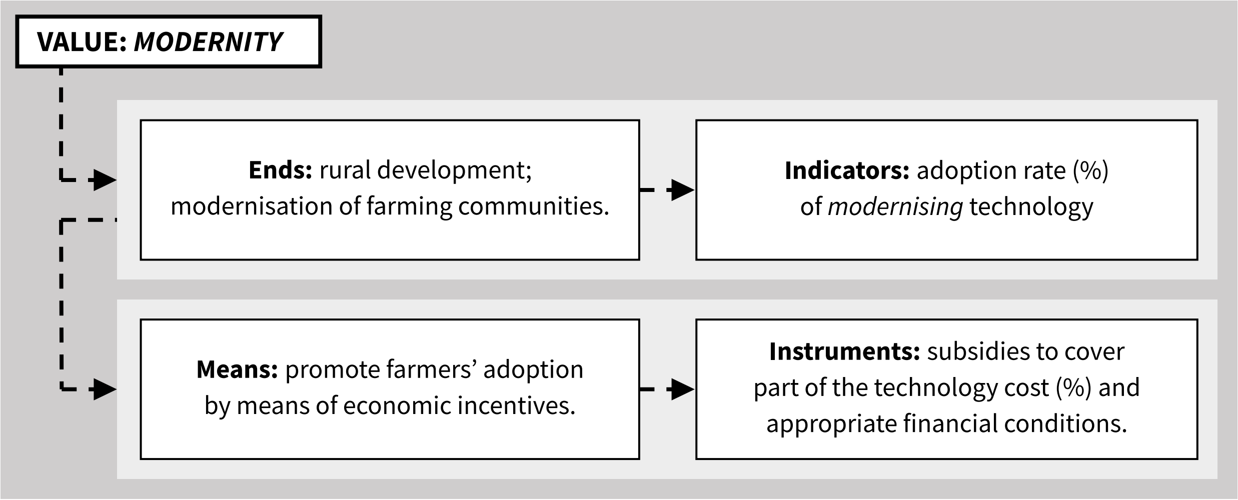


Figure 9.1: Policy-schema enacted in the farming communities' to imbue *modernity*

in Sec. 9.6). There are only farmer *policy-targets*, who are motivated by one value: *wealth*, which is measured by the profits resulting from the exploitation of their farms.The assumption of profit-driven farmers in regions where agro-industry is strongly market-oriented and water scarcity is high seems quite suitable, as the results show. Anyway, these encouraging results also point towards the need to consider other values, as other modernisation cases do not fit into this view.

Notice that this reflects a particular view of one of the collective problems in the water domain. It does not question the hydraulic paradigm of the twentieth century in Spain [148]: the main objective of the technological change is to increase the availability of water in order to eventually contribute to the growth of the economy of the region. Although during the dictatorship it was achieved by means of building large water infrastructure (enlarging the water input), the current intervention aims at increasing the water productivity by reducing the water losses of a given water input at a farm scale.

Table 9.1: Framework applied to the case study of technology adoption in farmer communities

Domain (model and data)		Water & economy, Crops & economy, Environment and climate, Farms, Innovation & economy (see Appendixes A and B)
Stakeholders	**Policy-maker**	User (e.g. National Agency)
	Policy-target	Farmers (connected by spatial proximity) (see Sec. 9.4.1)
Value model	**Policy-maker**	(user): *modernity* (innovation, rural development). The more adoption rate, the better.
	Policy-target	(farmers): *wealth*. The more profits from farming, the better. This can lead them to adopt new irrigation technologies (see Sec. 9.4.3).
Policy-schema		See Fig. 9.1

Given this purpose, this policy simulator system is tested on historical data from two Spanish irrigation communities during the period 1975–2010.

9.2 Introduction to the case study

The agricultural water cycle involves a large proportion of water resources, accounting for approximately 70 % of global water withdrawals [80]. In a foreseeable future of growing population and changing climatic conditions, where agriculture will have to increase its production some 50 % by 2050, those requirements are not likely to diminish [81]. These expectations shall produce increasing socio-economic stress on the water cycle unless sound policy is implemented and effective actions are taken.

There is widespread consensus that water could be better managed in agriculture, especially where traditional low-efficiency irrigation systems are commonly used [148]. Consequently, many stakeholders consider the modernisation of irrigation systems as an essential means for better water use in farmers communities.

The motivation behind this case study is to understand how modernisation takes place in such communities in order to provide input for policy-design, and it is therefore necessary to model different decision processes and profiles of stakeholders. In this case, Spanish "irrigation communities" (ICs) are modelled because they use a large share of water resources in Spain, because enough reliable data from some communities is available to set up a core working model that may be extended to other communities, and because they follow an innovation process that may be adapted to agricultural communities elsewhere.

The approach is to build a simple agent-based model (ABM) —a population of farmer agents and a socio-economic context that influences individual choices— to understand how the dispositions of individuals to modernise propagate in the community and end up in the actual adoption or rejection of the innovation (Sec. 9.3). In the model, individual decision-making is based on comparing the current farm performance against the expected return (see [139]), as well as taking social influence into account.

The model in this chapter explains a two-staged adoption of modern irrigation technology —collective and individual— with profit-driven individuals immersed in a social network where farm extension is a proxy for social influence (Sec. 9.4). Individual decisions (unknown from data) are modelled, from which the community adoption is simulated, fitting quite faithfully the actual adoption data. The calibration of the model was done with real data from two communities and indicated that, in these cases in which agriculture is market-oriented and water scarcity is high, favourable modernisation conditions arise from added-value crops, which are enabled by higher water allocations and greater irrigation efficiency (Sec. 9.5).

9.3 Background

Modernisation of irrigation systems is said to improve water use efficiency [148]. In traditional communities, water is distributed by means of open channels, and applied using field ditches. A drawback of such systems is their significant level of water losses. These contrast with modern systems with pressurised networks and sprinklers or drip pipes —whose adoption is known as "modernisation"—, that minimise water losses but consume more energy.

An irrigation community (IC) is a group of agents that share a water allocation right, whose main use is farm irrigation. In Spain, an IC has the power to plan and execute infrastructure projects, as long as an assembly constituted by all its members agrees. This procedure is based on direct voting, for which farmers hold a number of votes proportional to their farm extension. Actual modernisation of an IC (the object of this simulator) is a two-stage process: (i) the community upgrades the shared infrastructure only when supporters of modernisation gather enough votes to pass the project; and (ii) only then are farmers able to install drip or sprinkler systems in their own fields.

Diffusion of innovations is the process through which practices, ideas or products, are spread and adopted over time within a social system [205]. Diffusion of innovations in ICs has been studied using equation-based models [10]. However, this modelling approach does not reflect agent heterogeneity and does not reproduce farmers' social interactions explicitly.

Agent-based models can overcome these limitations and have been used to study adoption of innovations [63, 231, 275, 64]. Moreover, it allows to approach adoption-decision as a contingent phenomena, in which farmers may adopt a technology individually only after a prior collective agreement that emerges from the same decision-making units.

A line in the agricultural domain has taken a socio-hydrology approach [239], combining human behavioural models with biophysical models, to focus, in particular, on water issues. Becu *et al.* [27] presented the CATCHSCAPE model, whose central point was the impact of upstream agricultural water management on the farming activity downstream, for which they considered a spatial representation of the watershed explicitly. Hu *et al.* [117] explored the use of groundwater for irrigation, and introduced environmental, economic, social, and infrastructure aspects on agents' decision-making on groundwater pumping. They concluded that factors relevant in agents' decisions were temperature, precipitation, groundwater level and crop prices —variables that are reflected in the presented model, namely in water availability and climatologic conditions. However, that model lacked social interaction between agents, since the agents represented aggregation of multiple farmers.

Similar to this work, Holtz & Pahl-Wostl [113] used ABM to conduct historical research in a Spanish basin where irrigation agriculture was a key economic activity. However, they focused on land-use changes and groundwater use. They considered diffusion of irrigation technologies, multiple crop options, and cognitive biases as risk aversion and concluded that farmer models should consider more values than profit orientation (e.g. lifestyle).

Berger [31] developed a model for assessing different policy scenarios in the diffusion of innovations and resource use changes. The model was based on *cellular automata* modelling approach, since it was considered that spatial dimension was essential in agriculture models. Although some of the potentialities afforded by the spatial representation are not included in the present model (e.g. resource distribution, nutrient diffusion, return flows, etc.), they are easy to be implemented in future work, since the spatial dimension is already taken into account.

The usual approach of innovation-diffusion models base decision-making on utility functions [137], and also consider contagion processes [10, 113, 31]. Although other non-economic variables might be included in this function, it is assumed that most farmers (especially large-scale farmers) focus primarily on economic utility as the driving value [29]. For instance, [231] categorised farmers according to their motivation, distinguishing between large-scale farmers that pursued profit maximisation and small-scale farmers interested in conservation of rural lifestyle. In this case study, it is approached a class of ICs for which the profit-driven assumption is plausible due to the regional socioeconomic context. Nonetheless, it is recognised the need of considering additional values and interests (see [118]).

9.4 Model

The model represents an *irrigation community* (IC) as a set of farmer agents for which modernisation is achieved in two stages (Fig. 9.2). The first one accounts for the commitment of collective modernisation, in which each farmer evaluates whether it is worth modernising its farm, and then, all farmers need to reach a collective agreement based on an aggregation of their attitudes. In the subsequent

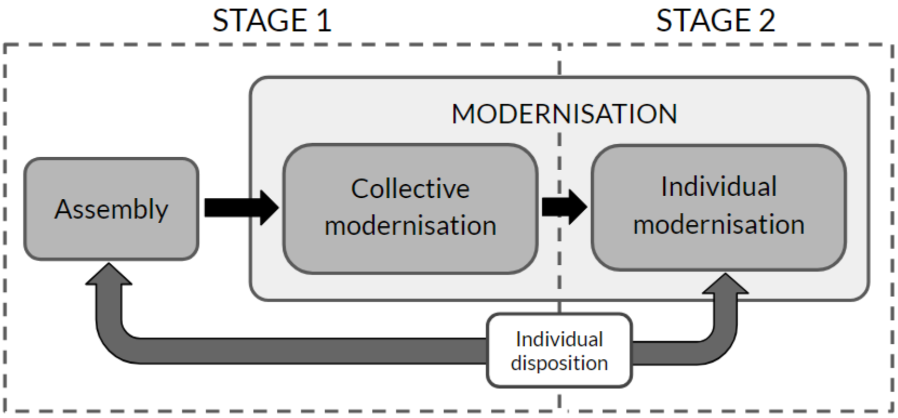


Figure 9.2: Modernisation of an irrigation community as a two-stage process

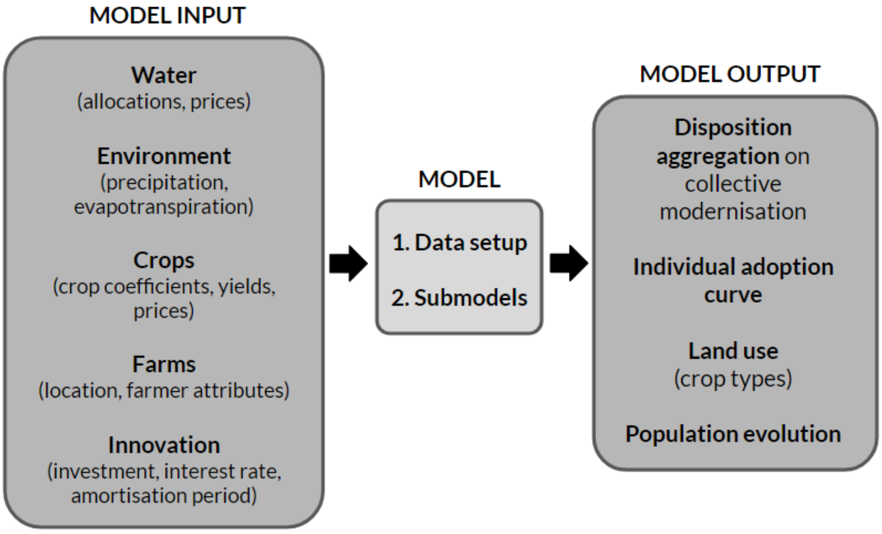


Figure 9.3: Components of the modernisation model

second stage, once the community agrees to modernise, individual farmers may choose to modernise their own farms. Information from experts and field data from irrigation communities are used to design and calibrate the model. Individual decisions (unknown from data) are modelled from which the community adoption is simulated (comparable with actual adoption data). Thus, the model can be used to gauge the proclivity of a given community to modernise and reveals external actions that may induce modernisation.

Five main types of data, drawn from several sources, are used as input for the model (Fig. 9.3, Appendix B). The individuals' disposition to modernise (Stage 1) and their actual adoption (Stage 2) is obtained as raw data, that then are aggregated to obtain the collective disposition and the adoption curve, respectively. *NetLogo* is used a simulation platform.

9.4.1 Entities and Assumptions

Farmers are the only kind of agent in the model. They are characterised by the following attributes: (i) *farm area*; (ii) *farm location*; (iii) *supply support* (capability to increase water supply by means of alternative sources, such as private wells); (iv) *crop-inertia* (reluctance to change crops in spite of potential profitability); (v) *risk-aversion*; (vi) *age*; (vii) *capital*; (viii) *past revenue* (economic outcome of the past agricultural season); and (ix) *disposition* towards modernisation. Attributes (i–vi) are based on real-data (see Appendix B) and attributes (vi–ix) evolve as results of the simulation.[1]

Social influence is a key element in the modernisation process, since farmers use information from other farmers to make decisions. Nonetheless, the replication of social networks in such agricultural communities is not a trivial task: social interaction between farmers may happen in multiple occasions (e.g. community assemblies, sporadic contact in town, familiar relationships, etc.). Indeed, this particular phenomena justifies more fieldwork.

Farmers' *social network* is based on two elements in this work: (i) spatial proximity and (ii) farm scale. Although spatial proximity (i.e. social distance) generation model may ignore some aspects that are relevant in social networks (for instance, that well-connected nodes are generally connected with each other), which is considered to be a plausible hypothesis for constituting a social network in a agricultural community. Hamill & Gilbert [103] proposed a similar mechanism to generate social networks, whose perception model is based on social circles (whose radius was labelled "social reach").

Homophily in the social network is introduced with the consideration of the farm scale. The assumption is that each farmer pays attention to similar neighbours, since their plots have similar farming features and their experiences are likely to apply. Scale is defined on farm extension, following the classification of [208] in Spanish communities: (i) *small scale* farmers are those that own between 0 and 20 hectares; (ii) *medium scale*, from 20 to 70 ha; and (iii) *large scale*, greater than 70 ha. Following experts' opinion and [29], it is assumed that all farmers may be linked, but *small scale* and *large scale* farmers do not influence each other.

Formally, farmer *i* and farmer *j* of the set of farmers *A* are connected $R(i, j)$ when the distance between them is below a maximum distance d_{max} and their farm scales are not *small* and *large* (Expr. 9.1). The social network N_i of the farmer *i* (the other farmers it is connected with) is given by Expr. 9.2. Note that social network N_i is defined at the initialisation step of the simulation and fix for the whole simulation.

$$\begin{gathered}(\forall i, j \in A)(R(i,j) \leftrightarrow (distance(i,j) \leqslant d_{max}) \wedge \\ \neg(scale(i) = \text{small} \wedge scale(j) = \text{large}) \wedge (i \neq j))\end{gathered} \tag{9.1}$$

$$\forall i \in A,\ N_i = \{j \in A : R(i,j)\} \tag{9.2}$$

Farmers have *dichotomous states* in three issues: (i) farming their lands and participating in the community (*active* or *inactive*); (ii) adoption of the innovation (*traditional* or *modernised*); and (iii) attitude towards modernisation (*willing to modernise* or *not willing to modernise*). A farmer *willing to modernise* will vote for collective modernisation in Stage 1 and will modernise its individual plot if possible in Stage 2. Notice that *willingness* is a state, whose transition is determined by the *disposition* as explained below (Eq. 9.6).

To start a simulation, all farmers start being *active*, *traditional*, *not willing to modernise*, and have the crop that best suits their initial water availability.

The spatial scale of the model represents the farming area of the community. The model simulates decades of activity through discrete one-year steps, although some submodels use one-month steps (for instance, the crop yield estimation considers one-month steps). All the submodels are explained in detail in Appendix A and data in Appendix B.

---

[1] To initialise the simulation, the *age* of farmers —attribute (vi)— is set with real data, but it also changes with time.

9.4.2 Process Overview

The following procedures are executed sequentially each time-step (one year) (See Appendix A for further details):

1. **Update allocation**: set water context variables, such as water allocation and water prices.
2. **Water availability**: determined by the water allocated to the farmer, expected rainfall, private water supply (e.g. wells), and the efficiency of the irrigation system (shared and individual irrigation infrastructure).
3. **Crop choice**: farmers choose one crop (from a list of options), aiming to maximise their income. For the sake of simplicity, crop choice depends only on water and crop-related variables. Crop water requirements are calculated as the reference evapotranspiration multiplied by a crop coefficient (see [16]).
4. **Update environment**: precipitation and evapotranspiration are updated and memorised by farmers to choose crops next year.
5. **Production**: *Revenue := Income - Cost*: *Income* is a function of crop market prices, farm area, and crop yield, which is adjusted according to water stress (see [245]). *Cost* adds up the costs of water, energy, farm inputs and amortisation of the individual's irrigation system.
6. **Activity?**: farmers decide whether to stop farming or not depending on the revenue and the capital they have.
7. **Modernisation? (only Stage 1)**: farmers' willingness to modernise is defined by two processes: individual decision-making and social influence.
8. **Assembly (only Stage 1)**: farmers vote for or against collective modernisation.
9. **Modernise (only Stage 2)**: those farmers who are *willing to modernise* their farms do so. There is no strict dependency of this decision and what the farmers voted in the assembly.
10. **Population evolution**: farmers that reach the retirement age trigger a generational replacement, and *inactive* farmers can become active again, transfer their land to a new farmer, or remain inactive.

9.4.3 Submodels

Stage 1: The community commits to modernise

Individual decision-making: It is based on opportunity costs of traditional systems. Taking into account that modernisation would have brought a higher volume of available water, farmers evaluate how it would have affected their economy. Revenue and expectation are calculated as the difference between *income* and *costs* (Eqs. 9.3 and 9.4). In this specific procedure, *past revenue* considers the traditional irrigation system (shared and individual), whereas *expectation* regards a modernised system. Collective adoption costs are introduced as water costs (i.e. tariff scheme), whereas individual costs are those of amortisation. These variables are different in each time-step *t*, depending on the decisions made by the farmer *i* and context variables.

$$PastRevenue_{i,t} := Income_{i,t} - Costs_{i,t} \tag{9.3}$$

$$Expectation_{i,t} := Income_{i,t}^{(modernised)} - Costs_{i,t}^{(modernised)} \tag{9.4}$$

Farmers must perceive that the utility of the *expectation* is greater than the *past revenue* (Eq. 9.5). Provided that this condition is met, *disposition* compares *expectation* and *past revenue* including a *risk-aversion* parameter γ_{S1} (Eq. 9.6). This parameter is greater than or equal to zero; when it is zero, the farmer has no risk aversion and will adopt the innovation as long as benefits are expected; on the contrary, a large value for this parameter means that the farmer is averse to changes and will adopt only if it entails significant profits. As the *disposition* ranges from zero to one, a negative result will be considered as zero.

$$Expectation_{i,t} \geqslant PastRevenue_{i,t} \tag{9.5}$$

$$Disposition_{i,t} := 1 - \gamma_{S1} \frac{PastRevenue_{i,t}}{Expectation_{i,t}} \tag{9.6}$$

Social influence: a farmer *i* receives an influence on its own disposition. The new disposition is is the weighted average of the *disposition* of its neighbours (i.e. the agents that constitute N_i), besides its own disposition prior the influence. The weights are determined proportionally to their *farm area* (namely, their *farm area* $Area_j$ divided by the total area $Area_T$ of the farmers in $\{N_i\} \cup \{i\}$) (Eq. 9.7).

$$Disposition_{i,t} := \sum_{j}^{\{N_i\}\cup\{i\}} \frac{Area_j}{Area_T} \cdot Disposition_{j,t} \tag{9.7}$$

State transition: Each year, the transition from *not willing to modernise* to *willing to modernise* is based on a probability P_t (Eq. 9.8). This transition is bidirectional, meaning that farmers can change their mind on whether to modernise or not. If the farmer was already *willing to modernise* (W), the complementary probability is used as the transition probability.

$$P_t[\,W \mid \neg W\,] := Disposition_t \qquad P_t[\,\neg W \mid W\,] := 1 - Disposition_t \tag{9.8}$$

Assembly: The community will commit to modernise only when more than half the votes are for modernising (that is, all farmers may cast their votes but only farmers who are *active* and *willing to modernise* vote in favour). The number of votes each farmer has is proportional to its farm area.

Stage 2: Individuals modernise their plots

Individual decision-making: The procedure is the same as in the Stage 1, except that the *risk-aversion* parameter γ_{S2} is now greater than before. According to expert's observations, in the first stage farmers do not have to commit to adopt; and, furthermore, collective costs are shared among all community members. However, in Stage 2 farmers will invest their own money in their own plot that they will eventually have to pay with their own income. Therefore, their commitment is more fragile.

Social influence: it is based on imitation. Farmers who have not adopted observe those farmers in their social network N_i who have already adopted and also are making a profit (Expr. 9.9) —thus, they use a temporary social network $\hat{N}_{i,t}$. Then, they calculate an expected income from the average revenue per unit of area (Eq. 9.10). More precisely, a farmer *i* at time-step *t*:

$$\hat{N}_{i,t} = \{j \in N_i : (Modernised(j) = \text{true}) \wedge (PastRevenue_{j,t} > 0)\} \tag{9.9}$$

$$Expectation_{i,t} := Area_i \cdot \left[\frac{1}{|\hat{N}_{i,t}|} \cdot \sum_{j}^{\hat{N}_{i,t}} \frac{PastRevenue_{j,t}}{Area_j} \right] \tag{9.10}$$

Expectation must meet the previous condition (Eq. 9.5). Afterwards, a transition probability is calculated as before (Eq. 9.6), but using a specific *imitation risk-aversion* parameter α_t. This parameter decreases over time because, as time passes, knowledge about the innovation (both internal and external to the community) is higher, and imitation is less risky. This is represented by a reduction parameter ψ (in %) that updates (linearly) the *imitation risk-aversion* at the initialisation α_{t0} (where t is the number of years since the initialisation at year t_0) (Eq. 9.11).

$$\alpha_t = \alpha_{t0} \cdot (1 - \psi t) \tag{9.11}$$

State transition: The procedure is the same as before (Eq. 9.8), except that it is not bidirectional, since farmers *willing to modernise* adopt the innovation immediately and never retract.

9.5 Simulation

Historical data from two Spanish irrigation communities is used: *Alhama de Murcia* and *Campo de Cartagena*. Both are located in the Segura river basin in SE Spain (Fig. 9.4). In this region, water scarcity is severe and the average annual precipitation is low (approximately 300 mm/year). Thus, ICs receive surface water transfers from the Tajo river and farmers often draw from groundwater resources. Climatological conditions (i.e. evapotranspiration and precipitation) are regional normal values during 1981–2010 and are assumed constant over time. Crop variables (i.e. crop coefficients, yields, and prices) are set using national sources [153, 202, 201, 154]. Following experts' opinions, individual modernisation is assumed to cost 3,500 eur/ha, to be paid in a 5-year time frame with a 2.5 % interest rate. Global efficiency of the irrigation system (taking into account distribution and application) is set as $\sim$40 % and $\sim$75 % for traditional and modernised systems, respectively. Finally, social network maximum distance is 4 km (see Fig. 9.5 as an example of spatial setup).

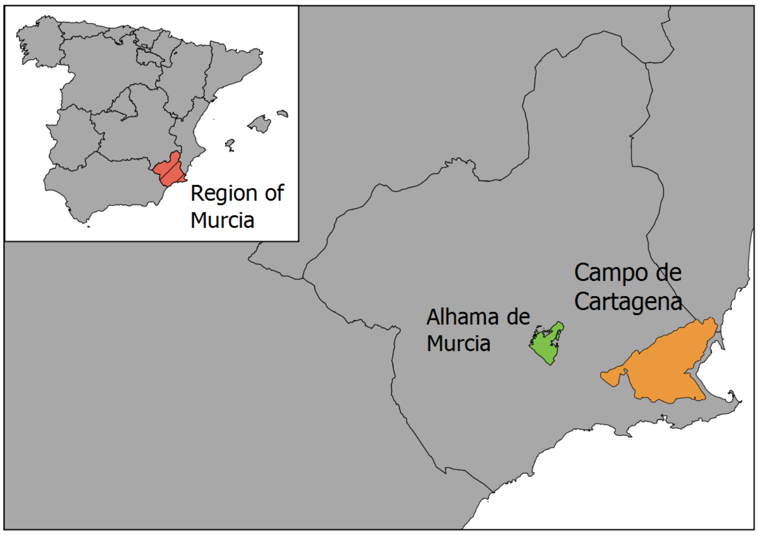


Figure 9.4: Location of studied irrigation communities in the Region of Murcia (Spain)

With regards to the actual modernisation process of the community as a whole, there is reliable data for both communities. The data set of the two communities is not complete, fortunately they mostly complement each other. Thus, given the similarities of the two ICs, the variables of each stage could be modelled drawing mainly on data from one community. Validating social processes of empirically-grounded models is difficult and has to cope with many uncertainties [182]. Replicative validation is considered at this point, comparing the simulation results with the real observations.

9.5.1 Stage 1: Collective modernisation

This stage is modelled mostly with data from *Alhama de Murcia*. This community has 5,906 hectares and 2,317 members. It started to operate in 1979 with a nominal allocation of 10,372,000 m^3/year, although since then, it has received only up to 50 %. In 1981, its General Assembly decided not to participate in the modernisation plan of the region. Nonetheless, it decided to modernise in 2000. Predominant crops in the area were citrus, grape, and other fruit trees and vegetables.

In order to simulate the *collective disposition*, data from 1979 to 2010 has been used. Water allocations are published [13], as well as water prices [229]. Crop options are a representative crop type in the aforementioned generic groups. Water fees are supposed to be 50 and 150 eur/ha/year, for traditional and modernised infrastructure, respectively. According to experts, collective modernisation costs are approximately 6,500 eur/ha with a payback period of 75 years. Data from a nearby community (namely *Campo de Cartagena*) has been used to define the individuals' attributes. In this stage, the risk-aversion parameter γ_{S1} in Eq. 9.6 is set empirically to 0.5 to fit observations, as done for Stage 2. Simulations involved 440 agents and 300 runs for each experiment.

Results indicate an increase of votes in 1985-1990 and in 1995–2005 —slightly higher— (Fig. 9.8). Figs. 9.6 and 9.7 show that they correspond with larger allocations and lower water prices, which enable new crop options that lead to added-value crops (Fig. 9.9). Crop-inertia plays a key role because it favours a change to crops with higher marginal water value (Fig. 9.8). During the 1985–1990 period, increased water supply leads to a mixture of two crops (Fig. 9.9): one with a higher water marginal value and lower water requirement (vegetables) and another with a lower water marginal value but larger value and larger water requirements as well (grape). In this period, most farmers that switch crops in the high-crop inertia scenario, grow grapes because they have larger water supplies. In the low inertia scenario, more farmers switch crops, including farmers with lower water supplies that allow them to move only to lower value crops (they switch to vegetables but cannot produce grapes). In the 1995–2005 period, inertia is less relevant because, due to the increased availability of water, producing grapes is feasible for almost all farmers.

9.5.2 Stage 2: Individual modernisation

The second stage of the model is tested using data from *Campo de Cartagena*. It comprises 37,433 hectares and their predominant crops were vegetables, citrus, and fruit trees. [10] collected data to build equation-based models of innovation diffusion for the period of 1975–2005. In that work, 360 farmers out of 3,237 were thoroughly characterised (e.g. *age*, *farm area*, *support-supply*, *crop-inertia*), and the water-related variables such as *water allocations* and *water prices* were given in detail.[2] In the absence of precise field data, representative crops in the region have been considered for the three predominant types (vegetables, citrus, and fruit trees). As suggested before, risk aversion in this stage is larger because farmers are now committing their own money. Thus, the *risk-aversion* parameter γ_{S2} in Eq. 9.6 is larger than the corresponding γ_{S1} in stage 1, namely, 1.085. The risk aversion when imitating α_{t0} is 1.970 with a reduction ψ of 2 % every year. These are equal for all farmers. In order to calibrate these parameters, different values for the *risk-aversion* parameters were tested against the actual adoption curve (see Figs. 9.12 and 9.13).

Adoption is triggered when innovators change crops and after that, the model shows that it spreads by imitation, regardless of crop patterns (Fig. 9.10). The adoption curve that results from the simulation resembles the typical logistic function, as expected, and fits the actual curve (Fig. 9.11).

---

[2] F. Alcón gave access to unpublished data from [10] that has been used to set these parameters.

Figure 9.5: Example of spatial setup of *Campo de Cartagena*. The triangles represent farmers, whose size is proportional to their farm area. Data on the location of farmers is provided at sector level, in which they are randomly distributed.

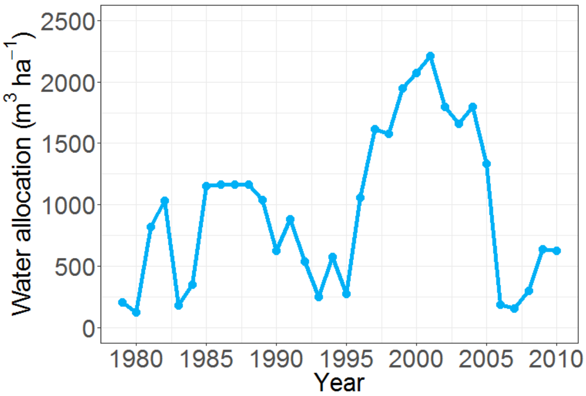


Figure 9.6: Water allocations for *Alhama de Murcia* during 1979–2010. They show prominent peaks in 1985–1990 and 1995–2005. *Campo de Cartagena* allocation distribution is similar, but volumes are approximately twice as large.

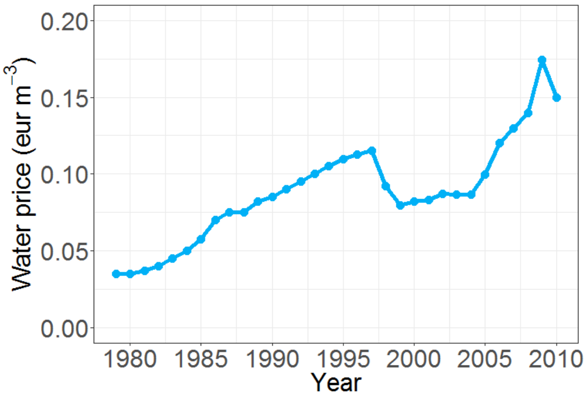


Figure 9.7: Water prices for Tajo-Segura transfers during 1979–2010. Prices increase from 1979, have a valley during 2003–2005, and then a steeper increase. These prices have been used for *Alhama de Murcia*.

9.6 Discussion

Simulations show reasonable correspondences with the studied cases in Spain. Significant increases in collective modernisation disposition are produced when farmers can leap to higher-value crops that have a greater water demand. These results are consistent with other works [148, 194].

The model is validated with real data. No information on individual farmers' behaviour and social

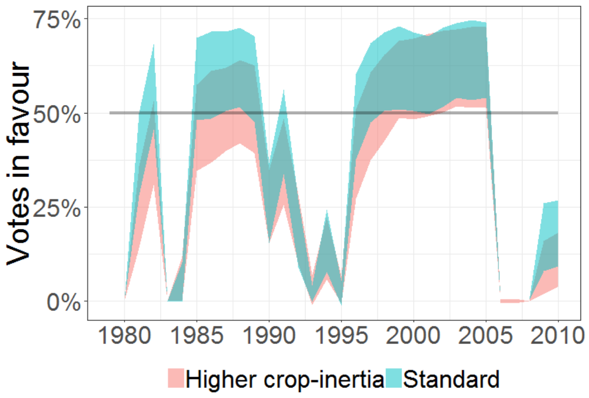


Figure 9.8: Simulated collective disposition in *Alhama de Murcia* during 1979–2010 under *standard* (real data) and *High* (twice the standard) crop-inertia.

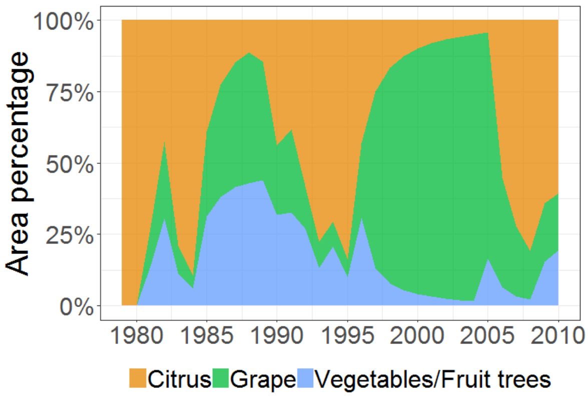


Figure 9.9: Simulated crop types in *Alhama de Murcia* during 1979–2010. Notice that crop pattern change with high water allocations and low water prices.

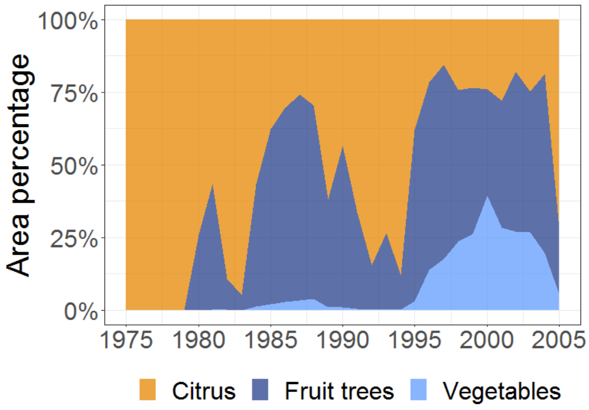


Figure 9.10: Simulated crop types in *Campo de Cartagena* during 1975–2005. Crop pattern change with high water allocations.

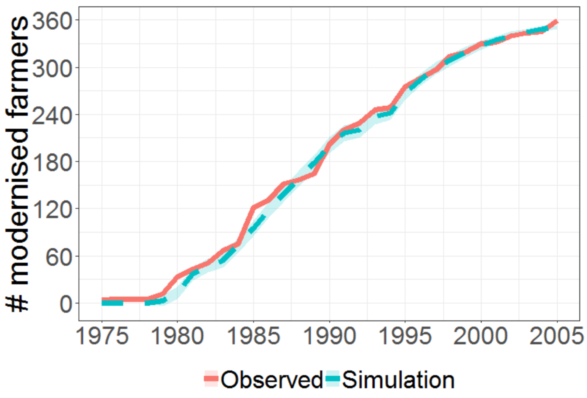


Figure 9.11: Simulated adoption curve in *Campo de Cartagena* during 1975–2005.

interactions was available beyond descriptive expert knowledge and agent attributes (e.g. farm area), but highly aggregated IC data (e.g. the adoption curve) was published. Comparing the actual results with the simulated output —that emerged from simulated individuals' behaviour and social influence models— indicated that they matched quite well.

The model presented in this chapter identifies when favourable conditions appear. In other words, what is relevant is the evolution of the collective disposition and its local maxima. Although the model has been calibrated to reflect credible voting results, additional empirical data about agents' resolutions is likely to improve calibration.

It can be deduced from the model that when water availability is too high, farmers do not perceive modernisation as advantageous, since their water demand would be completely satisfied and costs will increase unnecessarily. In other words, if farmers have as much water as they need to grow the highest water-demand crop —despite the water losses of the traditional system—, 'modernisation' is not attractive at all because it does not produce any extra profit. Although water losses are lower because of the innovation, they already have the volume they need for the chosen crop. In this case, "modernisation" does not lead to anything but higher costs.

Likewise, if water availability is low due to short allocations, farmers will not modernise since production increases do not compensate the costs or they cannot leap to higher-water-demand crops. Furthermore, uncertainty, and not only availability, affects decision-making [11]. This uncertainty could arise from environmental variables (rainfall resources and extreme weather events) or institutional reliability (whether the river basin authority or community can ensure the entire water allocation and water prices). Uncertainty has not been modelled explicitly in the model beyond the *risk-aversion* and the *imitation risk-aversion* parameters.

Varying *risk-aversion* supports intuitive results: the greater the risk aversion, the lower the adoption rate. Fig. 9.12 shows that a *risk-aversion* $\gamma_{S2} > 1$, (i.e. *statu quo* weighs more than future expectations) entails delayed adoption curves. Lower values for the parameter lead to a significant number of adoptions at the beginning, that are delayed by a period of low water allocations (Fig. 9.6) —since greater water availability due to modernisation does not result in higher profits—, and then recover to slowly enter in a stationary phase. Notice that the “exponential growth” in the adoption curve is still notable for $\gamma_{S2} > 1$, which reflects that the delaying factor is presumably the lack of early adopters that spread the innovation (who decide to adopt at the end of the period, when highest water allocations are provided) —which also explains the deviation in simulation results. This exercise reveals the considerable sensitivity of the parameter: for instance, valuing the *statu quo* only 0.85 in comparison to the *expectation* (see Eq. 9.6) leads to a quite different adoption curve.

Similarly, *imitation risk-aversion* α_{t0} shows higher sensitivity when $\alpha_{t0} < 2$ (Fig. 9.13). For greater values ($\alpha_{t0} > 2$), effects are reduced with respect to the base line, which may reveal that imitation, despite occurring, is infrequent. In other words, as the *imitation risk-aversion* parameter increases, the “social effects” are practically suppressed.

Since the shape of the estimated adoption curves follows a logistic curve as expected, the choice of the values for risk aversion parameters is guided by the extremes of the actual adoption curves. Better grounds for fixing risk aversion parameters should come from field data from these and other communities. However, as incidental support to this approach, notice that the shape of the estimated curves not only fits nicely the rest of the interval but also reproduce the “bumps” of the actual curves. Thus, for example, the model estimates that adoption is almost absent in periods of low water allocation (1981–1984 and 1987–1994 in Fig. 9.6), a phenomenon that is also visible in the real adoption curve and is easily observable in Fig. 9.13.

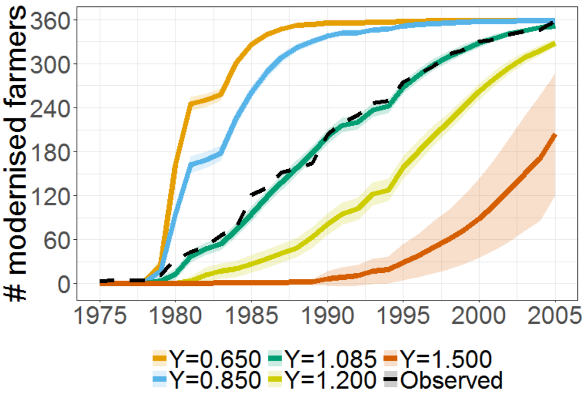


Figure 9.12: Simulated adoption curves in *Campo de Cartagena* using different values for *risk-aversion* γ_{S2}, compared with the observed data. Bands that surround curves represent the deviation in simulation outputs.

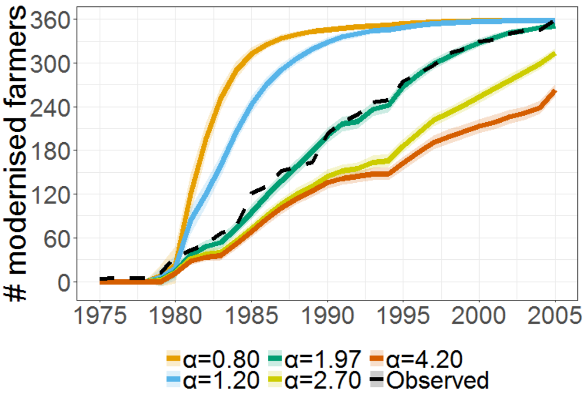


Figure 9.13: Simulated adoption curves in *Campo de Cartagena* using different values for *imitation risk-aversion* α_{t0}, compared with the observed data.

Modernisation affects the efficiency of the irrigation infrastructure, but it requires a proper management as well [61, 172]. Although these factors were not introduced in the model, it is expected that an efficient management would accelerate adoption and diffusion because of two reasons: (i) a poor management implies that not all potential benefits are entirely exploited, making modernisation less attractive to farmers; and (ii) untrained farmers are less successful, hence less likely to be imitated by other farmers. Since in this model imitators only perceive the consequences of the modernisation —i.e. revenue and not the new crop options— they fail to assess the total benefits. According to the simulations, this social influence is relevant for the second stage: without it, only ∼70 % of the farmers eventually adopt (Fig. 9.14). This also suggests that an appeal to profit-driven motivations is more likely to be successful when it is addressed to opinion leaders in those communities that are market-driven, able to adopt new crops and threatened with water limitations.

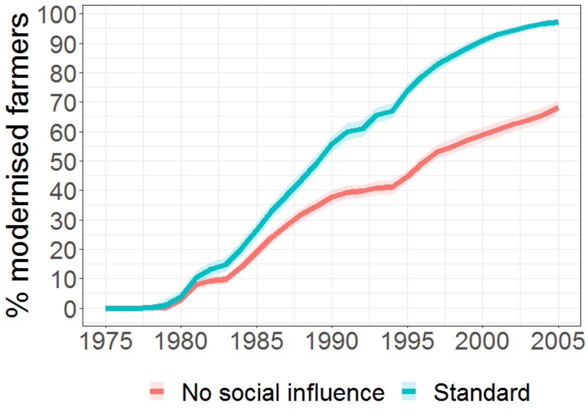


Figure 9.14: Simulated adoption curves in *Campo de Cartagena* for two scenarios: *standard* (actual adoption process) and *no social influence* (i.e. without social network).

Experts have proposed other conditions for modernisation that could be approached as extensions of the current model. Two are noteworthy: (1) [179] reported that some farmers were deceived: while modernisation was offered for free or strongly subsidised, in practice farmers had to pay for it, which hindered financial sustainability. In the presented model, it means that farmers used optimistic values which led to collective modernisation that would not have happened if they had not been misled. (2) In this model, collective decision is approached as a bottom-up aggregation of individual dispositions, a reasonable assumption for communities in which the social structure is horizontal. However, in some communities there is a small number of promoters of modernisation who try to influence other members to vote in its favour: politicians or community presidents promote modernisation in order to gain recognition or increase power (see [215]). In this case, only a few agents would decide (Eqs. 9.3–9.6), and the diffusion model (Eq. 9.7 and social network building) would consider other social characteristics to weigh the influence.

Edmonds [71] pointed out the problems of representing opinions as numerical measurements, and indicated the concerns when using such approach in *causal mechanisms* of the simulation model. Thus, representing the *disposition* to modernise by a simple numerical score, and use it in particular social influence functions, are strong assumptions. Modernisation may be motivated (or rejected) by combinations of other values besides economic profit: for instance, comfort, time savings due to automatisation, sense of progress, etc. In other words, decisions about modernisation arise, presumably, from a more complex model of rationality (which not only takes economic profit into consideration). This points for further research in how to work with values in computational models.

For these reasons, and although the results are encouraging, more field research is needed to better ground some assumptions in the model —for example, non-economic values (e.g. tradition, power, etc.), *risk-aversion* and *crop-inertia* parameters, social network generation— to apply it in communities whose socio-hydrological characteristics are different from the ones whose data has been used. Consequently, fieldwork should be done to further explore the farmers' values and their understandings —as pointed out by [118]—, since this will shape how the behaviour of the artificial farmer agents is modelled (see also [110, 219]).

Finally, it is convenient to remark that modernisation has been questioned from a perspective focused on environment conservation perspective [150, 148, 193]. It has been suggested that modernisation leads to farming intensification (e.g. extension of the farming area, double-cropping, or the adoption of crops with higher water demand), eventually increasing the use of water resources. One of the reasons, for instance, is that farmers are forced to intensify their production in order to compensate higher energy costs —phenomena that is aggravated because energy prices are continuously rising, while prices in agricultural markets follow a downward trend [150].

Although farming intensification in the model can only be produced by the adoption of crops with higher water-demand, the aforementioned phenomena is reflected in the simulations. Three scenarios have been simulated for *Campo de Cartagena*: (i) a *standard* scenario, where farmers adopt progressively (as they do in Fig. 9.11); (ii) a *traditional* scenario, in which individual farmers keep using traditional methods and do not adopt modern technology at any time; and (iii) a *modernised* scenario, in which farmers adopt the technology from the start of the simulation. As Fig. 9.15 shows, the three scenarios are almost identical and modernisation results in low water savings. What modernisation actually promotes is water productivity (Fig. 9.16) —which is, in fact, more sensitive to water allocation changes.[3] This indicates one of the trade-offs that public policy has to deal with, and different understandings of what 'efficiency' is in this particular domain (see also [193]). Therefore, it has been suggested that additional policy instruments are necessary if environment conservation (namely, water savings) is pursued, such as updating water allocations and water tariffs (see [148]).

9.7 Closing remarks

An agent-based model to explain modernisation of the irrigation system in farmer communities has been presented in this chapter. Modernisation is modelled as a contingent innovation-diffusion process: a first stage where the community establishes a collective agreement to modernise —based on individuals' dispositions— followed by an individual adoption decision. The model was built using historical data (1975–2010) from two Spanish irrigation communities where water is scarce and efficient irrigation (resulting in higher water availability) enables more profitable crop types that are commercially consolidated.

The model is predictive for modernisation under such particular conditions. It can be used by policy makers to foster modernisation in similar contexts, to incentivise or subsidise the emergence of the propitious conditions, or to either dismiss modernisation or identify (alternative) non-profit driven motivations.

Although results are promising, it is recognised that this model has some limitations. Conceptually speaking, *bounded rationality* of farmers should be further improved using more fieldwork. Also,

---

[3] Water productivity is calculated as the revenue (Eq. 9.3) divided by the gross water volume used from the water allocation (that is, before water losses). The gross water volume relative to the area (Fig. 9.15) may not reach the level of the water allocation (Fig. 9.19 in Appendix B) because (i) the allocation is distributed equally among the whole crop season, but (ii) the crop requirements are different for each month.

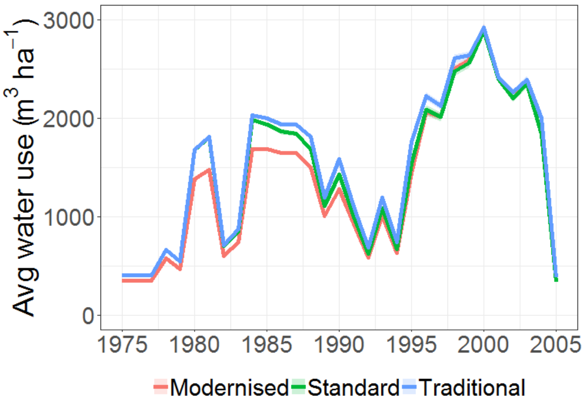


Figure 9.15: Simulated average water use (m^3/ha) in *Campo de Cartagena* from the water allocation (Fig. 9.19 in Appendix B) for different scenarios: *standard* (actual adoption process); *traditional* (farmers never adopt); *modernised* (farmers adopt from the start).

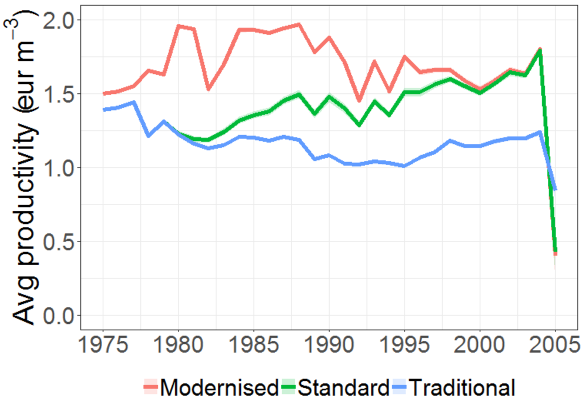


Figure 9.16: Simulated average water productivity (eur/m^3) in *Campo de Cartagena* for different scenarios: *standard* (actual adoption process); *traditional* (farmers never adopt); *modernised* (farmers adopt from the start).

some of the assumptions made to simplify the model or due to lack of data could be improved in further versions (e.g. there is no lag between crop-decision and crop-production, which may be inappropriate for tree-type crops; constant and current crop prices; etc.). Moreover, input from experts and available data may have driven the modelling process towards some consolidated worldviews about farmer communities.

Moreover, despite the fact that this model puts emphasis on water economy, the model could be extended to consider a broader approach for policy-making, namely taking the Food-Energy-Water nexus into account. As Cai *et al.* [46] pointed out, ABM is a promising modelling approach to consider that perspective. Besides, the case presented in this chapter is particularly interesting, as it combines agricultural economy, water governance, and the use of energy-intensive technologies [148, 150].

9.8 Appendix A: Submodels

This appendix complements Sec. 9.4. For the sake of simplicity, the subindex *t* (time-step) in all formulas, where time-step is one year, is omitted.

9.8.1 Water availability

Farmers estimate the expected volume of water available to irrigate for that year. This volume of water resources is given by:

- The *water allocation* of the community, $W_{allocation}$, for that particular year.
- The *supply support* S_i, which represents the share of the supplied water resources withdrawn from alternative sources like private wells. Therefore, this volume $W_{support,i}$ is calculated (Eq. 9.12):

$$W_{support,i} := \frac{S_i \cdot W_{allocation}}{1 - S_i} \tag{9.12}$$

- The *effective precipitation*, P_e, that is the fraction of precipitation that can be potentially used by crops, considering that there is a fraction that leaves the system (as runoff, for instance). It is assumed that only 75 % of the total precipitation P is effective precipitation.
- The irrigation infrastructure (constituted by the collective distribution and individual application systems), that determines the fraction of water that leaves the system (as leakage or evaporation) and cannot be used in crop evapotranspiration (Table 9.2).

The irrigation infrastructure determines the amount of water that is actually applied to the crop. A fraction of the total volume of water leaves the system due to system losses (i.e. leakage or evaporation). Thus, considering the efficiency of water distribution μ_d and the efficiency of water application $\mu_{a,i}$ (that depends on the farmer), the *water allocation* $W_{allocation}$ is reduced to (Eq. 9.13):

$$W'_{allocation,i} := W_{allocation} \cdot \mu_d \cdot \mu_{a,i} \tag{9.13}$$

Likewise, the support is reduced to (Eq. 9.14):

$$W'_{support,i} := W_{support,i} \cdot \mu_{a,i} \tag{9.14}$$

Table 9.2: Irrigation infrastructure efficiencies

System	**Collective, Distribution** (μ_d)	**Individual, Application** ($\mu_{a,i}$)
Traditional	0.75	0.55
Modernised	0.85	0.90

9.8.2 Water partition

It is assumed that the water allocation and supply support is distributed uniformly along the season —those months in which the crop requires to be irrigated, that is, $K_{c,j} \neq 0$ for month *j*—, although it is recognised that other considerations could be done at this point (for instance, that the volume is distributed equally along all the year regardless of the crop requirements).

Considering that the season comprises m months (i.e. the number of months with $K_c \neq 0$) (see Table 9.4 in Appendix B), then, the available water for each is (Eqs. 9.15 and 9.16):

$$W'_{allocation,i,m} := \frac{W'_{allocation,i}}{m} \tag{9.15}$$

$$W'_{support,i,m} := \frac{W'_{support,i}}{m} \tag{9.16}$$

9.8.3 Water irrigation

This procedure is done along the year for all the months. If the crop requires to be irrigated in month *j* ($K_{c,j} \neq 0$), then the farmer *i* uses all the water available for that month (Eq. 9.17):

$$W_{i,j} := W'_{allocation,i,m} + W'_{support,i,m} + P_{e,j} \tag{9.17}$$

The water requirement for crop *c* is calculated as the reference evapotranspiration $ET_{0,j}$ (which depends on climatological variables, such as temperature) multiplied by the crop coefficient $K_{c,j}$ [16] (Eq. 9.18):

$$ET_{c,j} = K_{c,j} \cdot ET_{0,j} \tag{9.18}$$

Then, the available water is compared to the required water to determine the volume of irrigation water *I*. Two cases are identified:

- **Case (a):** $W_{i,j} \geqslant ET_{c,j}$, which means that there is enough water to satisfy the crop water requirement. In this case, the farmer uses only the required volume (not all the available water) (Eq. 9.19):

  $$I_j := ET_{c,j} \tag{9.19}$$

- **Case (b):** $W_{i,j} < ET_{c,j}$, which means that the crop is under water deficit conditions. In this case, the farmer uses all the water that is available, although it does not satisfy the water requirement (Eq. 9.20):

  $$I_j := W_{i,j} \tag{9.20}$$

Notice that water used to irrigate has different sources (see Eq. 9.17). Accordingly, the farmer uses the different sources following a policy of minimising the cost of water. With this in mind, the first source to be used is precipitation, then the community allocation (Eq. 9.15), and finally the individual sources (Eq. 9.16), which are assumed to be more expensive than the other sources —mainly because they are groundwater resources and also not all the agents have access to them.

9.8.4 Crop yield estimation

Crop yield is estimated using the following relationship [245] (Eq. 9.21):

$$\left[1 - \frac{Y_a}{Y_{max}}\right] = K_y \left[1 - \frac{ET_a}{ET_{max}}\right] \tag{9.21}$$

where Y_{max} and Y_a are the maximum and actual yields, ET_{max} andET_a are the maximum and actual evapotranspiration, and K_y is a yield response factor, which determines the effect of a reduction in evapotranspiration on yield losses —that is, when the volume of irrigation does not satisfies the crop requirement (**Case (b)**)— (see Table 9.4 in Appendix B).

In the model, crop yield is addressed as follows:

1. First, the evapotranspiration ratio ET_a/ET_{max} is calculated for each month *j*. Using the variables as named in the previous submodels, it is computed as (Eq. 9.22):

   $$\xi_j := \frac{I_j}{ET_{c,j}} \tag{9.22}$$

   Notice that this ratio cannot be greater than one; if it is, it has to be set to 1.

2. Second, the yield ratio Y_a/Y_{max} is obtained following Eq. 9.21 considering the crop yield response factor $K_{y,c}$ (Eq. 9.23):

   $$\lambda_j := \left(\frac{Y_a}{Y_{max}}\right)_j = 1 - K_{y,c}(1 - \xi_j) \tag{9.23}$$

3. Third, the average of the yield ratio is estimated for the whole crop season (that comprises *m* months) (Eq. 9.24):

$$\bar{\lambda} = \frac{1}{m}\sum_{j}^{m} \lambda_j \tag{9.24}$$

Accordingly,

— In **Case (a)**, crop requirements are satisfied. Each additional water unit does not increase the crop yield, since the maximum production is reached. If farmers are considering to modernise their systems, the investment is less attractive, since the yield for that crop *c* does not change with more water.

— In **Case (b)**, the crop is under water deficit conditions. Each additional water unit increases the crop yield. Due to this fact, modernisation is more valuable, since it can lead to a greater production with the same water allocation (because less water is lost by leaks and evaporation).

9.8.5 Production

In this procedure, the economic output of farmers is computed.

First, the crop production is estimated (in tonnes), taking into account the actual crop yield (Eq. 9.24) and the characteristics of the crop (Table 9.4 in Appendix B) (Eq. 9.25). It can be used to calculate the income (in euros) (Eq. 9.26).

$$Production_{i,c} := \bar{\lambda} \cdot Y_{max,c} \cdot Area_i \tag{9.25}$$

$$Income_{i,c} := Production_{i,c} \cdot Price_c \tag{9.26}$$

Second, the farmer has to cover the costs of the activity. There are two main costs: (i) on the one hand, there are individual costs that are associated to the water application system and to farming as a productive activity; (ii) on the other hand, there are collective costs that are associated to the administration of the community and its collective water distribution infrastructure (i.e. investment, operation, maintenance, etc.). These costs are covered by water fees, which are established by the irrigation community, that are paid individually by all the farmers of the community. These fees are implemented by means of different water tariff schemes:

— A tariff scheme whose pricing is proportional to the farmers' irrigated area. This scheme is used by traditional communities (i.e. non-modernised).

— A tariff scheme whose pricing is a binary cost-allocation scheme: a fixed fee that is proportional to the irrigated area, plus a variable fee proportional to the volume of water used. This scheme is used by those communities that have modernised their collective distribution infrastructure. Some communities may base their water fees entirely on the variable term.

With this in mind, a farmer *i* has to consider multiple costs:

— Costs of water $C_{w,i}$ (Eq. 9.27 for traditional communities and Eq. 9.28 for modernised communities):

$$C_{w,i} := fee_{w,trad} \cdot Area_i \tag{9.27}$$

$$C_{w,i} := fee_{w,mod} \cdot Area_i + U_w \cdot Area_i \cdot \frac{I_{allocation,i}}{\mu_{a,i}} \tag{9.28}$$

where fee_w is the fixed water fees for the community (which is different depending on the community has its collective infrastructure modernised or not) in eur/ha; U_w is the variable water fee (i.e. water price) in eur/m^3; $I_{allocation,i}$ is the total volume of irrigation (i.e. all the year) in m^3/ha (see Eqs. 9.19 and 9.20) whose source is the water allocation of the community, and which is divided by the efficiency of application $\mu_{a,i}$ to obtain the actual registered volume at the entrance of the farm.

— Costs of operation and maintenance of the farm $C_{O\&M,i}$ (Eq. 9.29):

$$C_{O\&M,i} := U_{O\&M} \cdot Area_i \tag{9.29}$$

where $U_{O\&M}$ are the unitary costs of operation and maintenance in eur/ha (that are assumed to be 7,000 eur/ha in the simulation). This cost adds up the costs of farm inputs, fuel, etc.

— Costs of private supply $C_{p,i}$ (Eq. 9.30):

$$C_{p,i} := U_p \cdot Area_i \cdot \frac{I_{support,i}}{\mu_{a,i}} \tag{9.30}$$

where U_p is the unitary cost of the private water sources in eur/m^3; $I_{support,i}$ is the total volume of irrigation (i.e. all the year) in m^3/ha (see Eqs. 9.19 and 9.20) whose source is private, and which is divided by the efficiency of application $\mu_{a,i}$ to obtain the actual volume taking into consideration application losses. In the simulation, U_p is assumed to be twice the $fee_{w,mod}$.

— Costs of amortisation $C_{a,i}$ (to replace the individual application system).

The total cost is calculated as (Eq. 9.31):

$$C_A := Inv_i + \sum_{t=1}^{T} Inv_i \cdot (1 - \frac{t}{T}) \cdot r \tag{9.31}$$

where Inv_i is the total investment given by the unitary cost of the application system U_t in eur/ha (which is assumed to be constant —that is, no scale economies) and the farm area $Area_i$; r is the interest rate; and T is the lifespan of the technology. For traditional application systems (i.e flood irrigation or field ditches), U_t is assumed to be 600 eur/ha, while it, for modernised application systems (i.e. drip or sprinkler irrigation), is 3,500 eur/ha. The interest rate r is set to 2.5 %, and the lifespan is 15 years for both systems.

With this in mind, and considering equal payments along the lifespan, the amortisation cost is (Eq. 9.32):

$$C_{a,i} := \frac{C_A}{T} \tag{9.32}$$

Then, the total costs for a farmer *i* are calculated as (Eq. 9.33):

$$Costs_i := C_{w,i} + C_{O\&M,i} + C_{p,i} + C_{a,i} \tag{9.33}$$

Finally, the revenue is computed as the difference between the income (see Eq. 9.26) and the costs (see Eq. 9.33) (Eq. 9.34):

$$Revenue_i := Income_{i,c} - Costs_i \tag{9.34}$$

9.8.6 Crop choice

In this procedure, farmers choose the crop type for their farms in order to maximise their revenue. Nonetheless, they may be reluctant to change crops in spite of potential profitability due to risk aversion, specialisation, comfort, etc. All these factosr are reflected by *crop-inertia* ϕ. It is implemented as (Eq. 9.35):

$$P_i[change\,crop] = 1 - \phi_i \tag{9.35}$$

If this event is unsuccessful, the farmer will grow the same crop as the previous year. On the contrary, if the event is successful, the farmer will explore the multiple crop options (see, for instance, Table 9.4 in Appendix B). This consideration may be seen, somehow, as path-dependence on farmers' decision-making (as their decision depends, to some degree on their current land-use).

Taking into account the actual irrigation system (i.e. collective and individual), and using the previous submodels (namely, **water availability**, **water partition**, **water irrigation**, **crop yield estimation** and **production** submodels), it is possible to estimate the actual revenue for each crop option. The chosen crop is the crop option that maximises their revenue.

9.8.7 Actual production

Assuming that a farmer has chosen a crop, and using the previous submodels (namely, **water availability**, **water partition**, **water irrigation**, **crop yield estimation**, **production**, and **crop choice** submodels), the actual revenue of the farmer can be computed, which determines the variations of their *capital*. The farmer then memorises this revenue as *past revenue* (Eq. 9.3).

9.8.8 Modernisation?

Note: as this submodel has been explained in Sec. 3, this subsection will only provide some further details, and will not reproduce again the entire submodel.

In Stage 1, farmers evaluate how modernisation will impact on their economy. To do so, they compare the results between the current irrigation system (that is, traditional collective and individual systems) against a modernised irrigation system (both collective and individual systems). Generally, modernisation may increase the water availability, which can open new crop options that lead to greater income. However, costs also increase, mainly due to the investment that has to be made to install such systems.

In specific terms, farmer compare the revenue (Eq. 9.34) that is produced using the current traditional system (Eq. 9.3) with the one that is produced using a modernised system. Namely, an *expectation* is generated (Eq. 9.4) making use of the previous submodels (i.e. **water availability**, **water partition**, **water irrigation**, **crop yield estimation** and **production** submodels) assuming that a modernised irrigation system is being used. This *expectation* is then compared to the *past revenue* (Eq. 9.5).

Notice that the individual modernisation costs (Eqs. 9.31 and 9.32) may take into account different payback periods (for instance, in the simulation, farmers evaluate modernisation as they would have to pay the investment in 5 years).

In Stage 2, farmers use the same process of decision-making, but the conditions have changed slightly (the *risk-aversion* parameter, the collective infrastructure, and the social influence process) (see Sec. 3). Once the community decides to modernise their collective infrastructure (that is, the community goes from the Stage 1 to the Stage 2), farmers are *not willing to modernise* their individual application systems any more, and have to make the decision again with the new conditions.

9.8.9 Assembly

In Stage 1, the community holds an assembly to pass the proposal of modernising the collective infrastructure. The system is based on direct voting, in which each member of the community has a specific number of votes proportional to the area of its farm.

The community will commit to modernise only when more than half of the votes are for modernising. In practice, that means that more than 51 % of the area of the community is subject to be modernised, as the farmers who own that extension are *willing to modernise*. Besides, a farmer cannot have 51 % of the votes by itself (that is, at least two farmers have to vote in favour).

For the sake of simplicity, only farmers who are *active* and *willing to modernise* vote in favour.

9.8.10 Activity?

Each year, farmers decide whether to quit or not depending on the revenue from the previous seasons and the capital they have.

If the actual revenue is negative (that is, the farmer is losing money), the farmer is more prone to quit: facing three consecutive years of negative results will force the farmer to become *inactive* and leave their farm unproductive. Likewise, if the individual capital drops to below than 0 eur, then the farmer will become *inactive* too.

9.8.11 Population evolution

Each year, farmers' *age* increases by one. This can lead them to reach the retirement age and trigger a generational replacement.

When farmers reach the retirement age, they can either (a1) be replaced by a new farmer or (b1) retire and leave the farm unproductive (for instance, because they do not manage to find a successor). In this model, the retirement age has been set to 80 years old; the probability of event (b1) is 0.60; and the probability of event (a1) is the complementary.

In case of event (a1), the new farmer is between 18 and 45 years old, and *not willing to modernise* (that is, the decision process has to be made again). Notice that, if the individual system of the farm has been already modernised, the new farmer has no decision to make in that matter.

Moreover, *inactive* farmers can (a2) remain inactive, (b2) become *active* again, or (c2) transfer the land to a new farmer.

It has been assumed that the probability of event (a2) is 0.80, and the probability of both events (b2) and (c2) is the complementary. In this latter case, the probability of event (c2) is 0.05. As before, the new farmer is between 18 and 45 years old, and *not willing to modernise* (that is, newcomers have to make the decision process again).

9.9 Appendix B: Input data

A summary of the input data for the simulations can be found in Table 9.3.

9.9.1 Crop options and characteristics

According to Alcón [10], in *Campo de Cartagena*, predominant crops are vegetables, citrus, and fruit trees —at that time, they accounted for 51 %, 35 % and 8 %, respectively. In *Alhama de Murcia*, predominant crops included citrus (45%), vineyards (35 %), and vegetables and fruit trees (20 %) in 2018 [14].

Table 9.4 shows the crop options that have been considered in the simulation. Crop characteristics (i.e. evapotranspiration factors, maximum yield, prices, and yield response to water stress) were collected from diverse sources [245, 16, 153, 202, 202, 154]. In the absence of precise field data, representative crops in the region have been chosen for each predominant crop type: namely, orange (*Citrus*),

Table 9.3: Summary of input data

Variable		Input	Sources
Crops	Crop factors K_c	See Table 9.4	[16, 154]
	Yield water response K_Y	See Table 9.4	[245]
	Maximum yield Y_{max}	See Table 9.4	[153, 202, 201]
	Crop prices $Price_c$	See Table 9.4	[153, 202, 201]
Climate	Effective precipitation P_e	See Table 9.5	AEMET
	Evapotranspiration ET_0	See Table 9.5	AEMET
Water economy	Water prices U_w	See Figs. 9.18, 9.20	[10, 229, 13]
	Water allocations $W_{allocation}$	See Figs. 9.17, 9.19	[10, 13]
	Water fees $fee_{w,trad}$ (Stage 1)	50 eur/ha	Set by experts
	Water fees $fee_{w,mod}$ (Stage 1)	150 eur/ha	Set by experts
Farmers	Age	See Table 9.6	[10], INE
	Farm area $Area_i$	See Tables 9.7, 9.8	[10, 14]
	Supply-support S_i	See Tables 9.9, 9.10	[10]
	Crop-inertia ϕ_i	See Tables 9.11, 9.12	[10]
	Risk aversion γ_{S1} (Stage 1)	0.5	Set by authors
	Risk aversion γ_{S2} (Stage 2)	1.085	Set by authors
	Imitation risk aversion α_{t0} (Stage 2)	1.970	Set by authors
	Imitation risk aversion reduction ψ (Stage 2)	2 %	Set by authors
Individual modernisation	Unitary cost U_t	3,500 eur/ha	Set by experts
	Interest rate r	2.5 %	Set by experts
	Payback period T	5 years	Set by experts

lettuce (*Vegetable*), peach (*Fruit-tree*), and lettuce + watermelon (*Vegetable2*). *Grape* is only available for *Alhama de Murcia*.

Table 9.4: Crop options and crop characteristics used in the simulation

Crop	**K_C**												**Y_{max}**	**K_Y**	**Price**
	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**M12**	T/ha	(–)	eur/T
Null	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Citrus	0.5	0.5	0.6	0.6	0.6	0.7	0.7	0.7	0.7	0.6	0.5	0.5	40	0.85	390
Fruit-tree	0	0.4	0.6	0.8	1.0	1.0	0.7	0.2	0	0	0	0	30	1.1	550
Vegetable	0.4	0	0	0	0	0	0	0	0	0.6	0.9	0.9	30	1.15	150
Vegetable2	0.4	0	0.2	0.3	0.9	0.9	0.6	0	0	0.6	0.9	0.9	55	1.1	275
Grape	0	0	0	0.5	0.6	0.7	0.7	0.3	0	0	0	0	26	0.85	765

9.9.2 Climatological conditions

Climatological conditions (i.e. reference evapotranspiration ET_0 and precipitation P) are regional normal values per month during 1981–2010.[4] In the absence of precise historical meteorological data, they are assumed constant over time (that is, each year values are repeated).

Effective precipitation (P_e) is the fraction of precipitation that can be potentially used by crops, considering that there is a fraction that leaves the system (as runoff, for instance). It is assumed that only 75 % of the total precipitation is effective precipitation (although it is recognised that more sophisticated models are available to estimate this fraction). These variables are usually given in millimetres or L/m^2; they have been converted to m^3/ha to facilitate the operation with crop-related models.

Table 9.5 shows the climatological conditions —for each month of the year— that have been used in the simulation.

Table 9.5: Climatological variables used in the simulation (in m^3/ha)

	M1	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**	**M9**	**M10**	**M11**	**M12**
ET_0	561	638	880	1153	1428	1740	1791	1625	1200	865	590	497
P_e	192	145	163	242	151	23	13	17	140	343	229	247

9.9.3 Water prices and water allocations

For *Alhama de Murcia*, water fees are supposed to be 50 eur/ha/year for the traditional collective infrastructure (see Eq. 9.27). For the modernised collective infrastructure, water fees are supposed to be 150 eur/ha/year (fixed fee), to which the variable fee that depends on the volume used to irrigate and the *water price* has to be added (see Eq. 9.28). *Water allocation* and *water prices* can be seen in Figs. 9.17 and 9.18. This data was obtained from [13, 229]. [5] [6]

---

[4] `http://www.aemet.es/es/serviciosclimaticos/datosclimatologicos/valoresclimatologicos?l=7031&k=mur`

[5] `https://www.cralhama.org/el-trasvase-tajo-segura-y-los-regadios`

[6] `http://www.scrats.es/tarifas-vigentes.html`

For *Campo de Cartagena*, the tariff scheme is based entirely on the volumetric term (see Eq. 9.28). *Water allocation* and *water prices* can be seen in Figs. 9.19 and 9.20. This data was obtained from [10]. Notice that, for both communities water allocations follow a similar distribution, although *Campo de Cartagena*'s is almost twice as large.

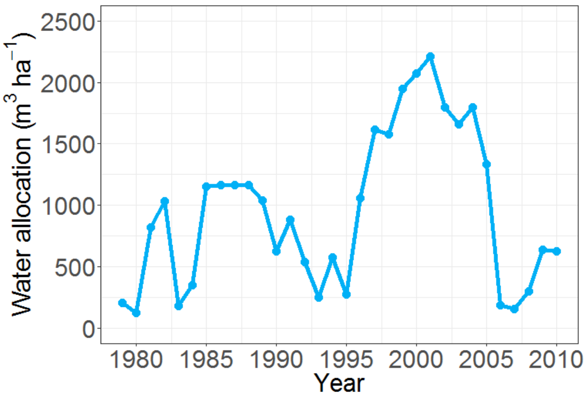


Figure 9.17: Water allocations for *Alhama de Murcia* during 1979-2010

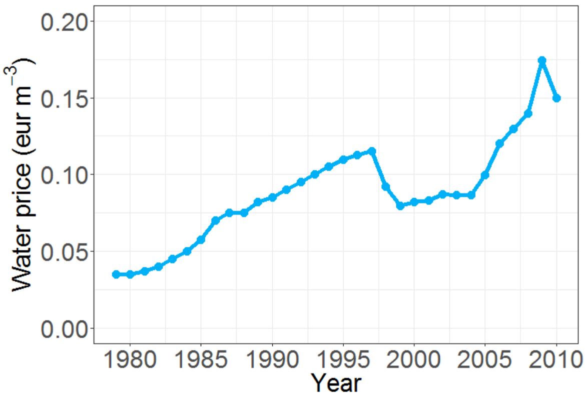


Figure 9.18: Water prices for Tajo-Segura transfers during 1979-2010 used for the simulation of *Alhama de Murcia*

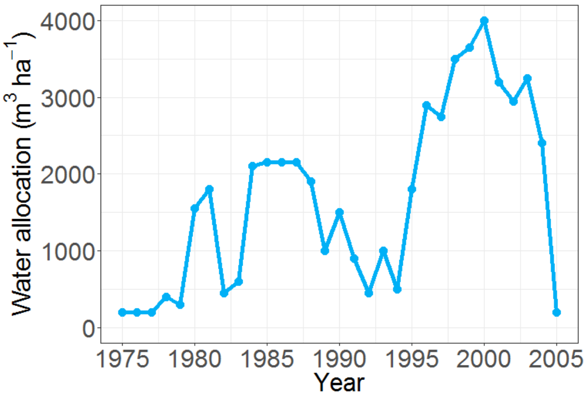


Figure 9.19: Water allocations for *Campo de Cartagena* during 1975-2005

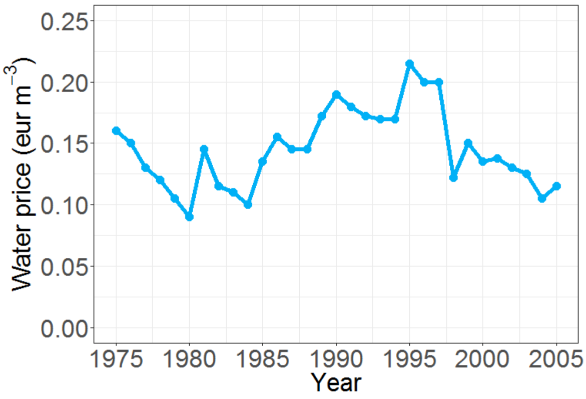


Figure 9.20: Water prices for *Campo de Cartagena* during 1975-2005

9.9.4 Farmers characterisation

F. Alcón made available unpublished data from [10] that has been used to characterise farmers in the model. Alcón conducted field surveys and collected data about farmers in *Campo de Cartagena*. Consequently, access to individual agent's data for variables such as *age* and *farm area* as available, but in this appendix only aggregated data is shown due to privacy concerns.

In the absence of data for *Alhama de Murcia*, data from [10] was used to characterise farmers in that community. For this purpose, the values for farmers variables were randomised using normal distributions whose mean and deviation were obtained from the sample of *Campo de Cartagena*.

Age

Table 9.6 shows the characterisation by age of farmers. In *Campo de Cartagena*, *age* was characterised according to [10]. In *Alhama de Murcia*, age was set following the statistical distribution from the Spanish agricultural census in 2009.[7] In this latter case, as only aggregated data was available, a random value in the range was generated following a uniform distribution.

Table 9.6: Age distributions used in the simulation

Age range	***Alhama de Murcia* proportion** (%)	***Campo de Cartagena* proportion** (%)
<25	0.3	0.3
25–34	4.5	12.5
35–44	13.6	28.7
45–54	22.3	19.8
54–64	26.3	22.8
>65	33.0	15.9

Farm area

Table 9.7 shows the distribution of farmers by farm area in *Campo de Cartagena*, which was set according to [10].

In *Alhama de Murcia*, field data was available, but it does not allow a precise characterisation of farmers [14].[8] They provide the total number of farms, and their distribution by areas using broad ranges —which are followed in Table 9.8—, as well as the total area for each category. Noteworthy, most of the farms are below 1 ha. Notice that these data are referred to farms, not to farmers. For this reason, this data needs some conversion.

The same proportion of farms in each category is used to determine the proportion of farmers in that same category. Then, using the total area of that category, it is obtained an average area per farmer. Following the same categories, and using the data from *Campo de Cartagena*, a deviation for each category is obtained. With this average area and area deviation in mind, farmers' farm area were randomised using normal distributions (assuming a minimum and maximum areas of 0.1 and 500 ha, respectively). Table 9.8 shows the distribution of farmers by farm area in *Alhama de Murcia*.

Supply-support

For *Campo de Cartagena*, *supply-support* was inferred from the data collected by [10] with the question: "*Water used on the farm: Origin (i.e. surface, groundwater), Source, and Use share (%)*". With this in mind, *supply-support* was assumed to be the value of the use share. Table 9.9 shows the distribution of farmers by *supply-support* in *Campo de Cartagena*.

---

[7] `https://www.ine.es/jaxi/Tabla.htm?path=/t01/p042/a2009/prov00/l0/&file=1101.px&L=0`

[8] `https://www.cralhama.org/distribucion-de-cultivos-propiedad-y-sistema-de-riego/`

Table 9.7: Farmers distribution by farm area in *Campo de Cartagena*

Category	**Proportion** (%)
< 1 ha	3.9
1–2 ha	4.2
2–5 ha	10.0
5–10 ha	13.1
10–20 ha	22.3
20–30 ha	11.7
30–50 ha	17.5
50–70 ha	6.7
70–100 ha	3.3
⩾ 100 ha	7.2

Table 9.8: Farmers distribution by farm area in *Alhama de Murcia*

Category	**Proportion** (%)	**Average area per farmer** (ha)	**Area deviation** (ha)
< 1 ha	67.1	0.572	0.250
1–5 ha	28.4	2.612	0.940
5–10 ha	2.6	8.703	1.100
> 10 ha	1.8	46.310	83.200

For *Alhama de Murcia*, this data was randomised using normal distributions whose mean and deviation were obtained from the sample of *Campo de Cartagena*, taking into account the *scale* of farmers, assuming a minimum and maximum value of 0 and 100, respectively. It is recognised that this method is questionable, as the probability densities of the values over 100 and below 0 are given to 100 and 0, respectively. Table 9.10 shows the characterisation of *supply-support* for *Alhama de Murcia*.

Crop-inertia

For *Campo de Cartagena*, *crop-inertia* was inferred from the data collected by [10]. Namely, the question was; "*State your degree of compliance with the following statement, scoring from 0 to 10: Would you grow a very risky product that can generate a lot of profit?*". Using the answer given by farmers (that is referred to as *risk-affinity* score), *crop-inertia* was assumed to be $\phi_i = 1 - score$. Table 9.11 shows the distribution of farmers by risk-affinity score in *Campo de Cartagena*.

For *Alhama de Murcia*, this score was obtained using normal distributions whose mean and deviation

Table 9.9: Farmers distribution by *supply-support* in *Campo de Cartagena*

Supply-support (%)	**Proportion** (%)
= 0	23.7
0–10	13.1
10–25	20.3
25–50	38.2
50–75	3.9
75–100	0.8

Table 9.10: *Supply-support* characterisation in *Alhama de Murcia*

Farmer scale	**Average *supply-support*** (%)	***Supply-support* deviation** (%)
Small	21.5	19.8
Medium	28.5	18.8
Large	33.9	19.5

Table 9.11: Farmers distribution by risk-affinity score for calculating *crop-inertia* in *Campo de Cartagena*

Risk affinity score (0–10)	**Proportion** (%)
Equal to 0	8.6
1–5	36.2
5–9	44.8
Equal to 10	8.9
DK/NA/REF	1.4

were obtained from the sample of *Campo de Cartagena*, taking into account the *scale* of farmers, and assuming a minimum and maximum value of 0 and 10, respectively. As in the previous case, *crop-inertia* was calculated as $\phi_i = 1 - score$. Table 9.12 shows the characterisation of *risk-affinity* score for *Alhama de Murcia*.

Table 9.12: *Risk-affinity* characterisation for calculating *crop-inertia* in *Alhama de Murcia*

Farmer scale	Average risk-affinity score	Score deviation
Small	5.1	3.2
Medium	6.1	2.7
Large	6.9	3.0

Farm location

Netlogo was combined with the *GIS* package to input the geographical location of farmers in the community.

Alcón [10] provided some rough geographical location of farmers in *Campo de Cartagena*. The surveyed farmers reported the sector of the community where their farm was located. In the simulation, the exact position of the farmers was randomised within the sector they reported (see Figs. 9.21 and 9.22). [9]

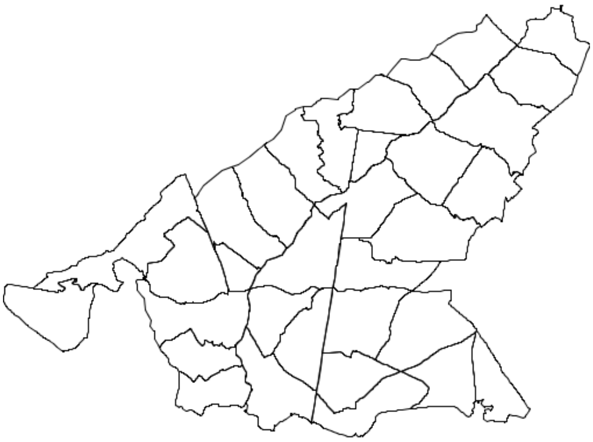


Figure 9.21: Geographical representation of *Campo de Cartagena*, divided by sectors.

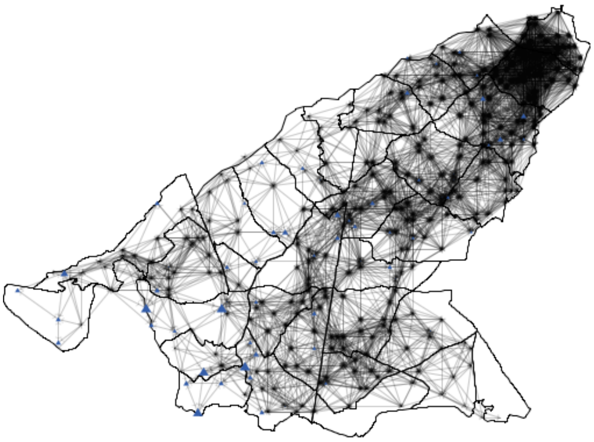


Figure 9.22: Randomly generated distribution of farmers in *Campo de Cartagena*.

This information was not available for *Alhama de Murcia*. Consequently, the exact location of farmers was randomly generated (Figs. 9.23 and 9.24).[10] [11]

---

[9] https://www.crcc.es/informacion-general/documentos-y-planos/

[10] https://www.cralhama.org/zona-regable/

[11] https://www.chsegura.es/chs/cuenca/resumendedatosbasicos/cartografia/descargas/

Figure 9.23: Geographical representation of *Alhama de Murcia*.

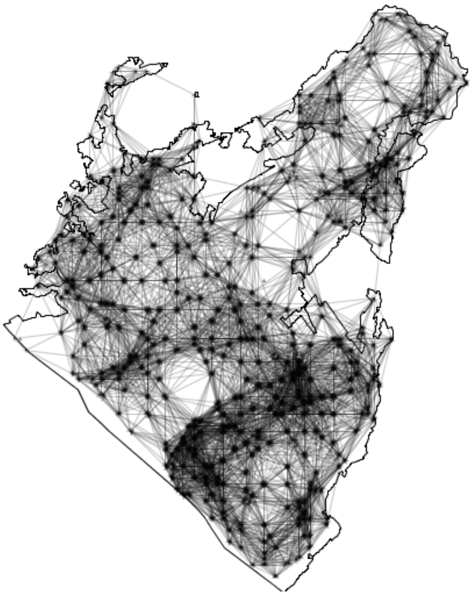


Figure 9.24: Randomly generated distribution of farmers in *Alhama de Murcia*.

Chapter 10

Case study: Modelling policy-shift advocacy in the urban water domain

The purpose of this chapter is to illustrate a generic ABM framework to test policies and social acceptance using values. The model presented in this chapter focuses, in particular, on the water public service in urban settings. On the one hand, *households* make use of the water services for their basic needs, but they want the service to be managed according to their values. On the other hand, *policy-influencers* demand political measures if they consider that the world-state is not aligned with their values. Thus, policy effects have an impact on social acceptance that may lead to policy shifts, thus altering the outcome.

10.1 Case study for the conceptual framework

Picture a neighbourhood of a city: each household houses a family with a certain income level, water needs, and conservation practices. There is a water utility company that supplies water, a public service that is supported by a fee. Citizens assess the service they get and may at some point want to have better conditions. Their satisfaction depends on what they believe is important (i.e. values), and they may identify some ways of intervening politically in order to increase their level of satisfaction. However, this is based not only on the service they receive directly, but also on some features that affect the society as a whole. Likewise, there are other stakeholders that assess the state of affairs with respect to their own values and may promote adjustments in the way water is being governed. As a consequence, there is an interaction between political actors and citizens that stimulates political action and policy shifts.

This case study is quite more complex than the case presented in the previous chapter. It serves to illustrate how complicated it turns the modelling process when addressing a case where multiple stakeholders with diverse values and political beliefs are involved. Besides, all this has to be build around the *policy-targets* —in this case, households—, whose behaviour is practically the main issue of the model, but whose “activity” is not as defined as in the case of farmers —for which water is clearly an economic input. Families behave differently, in accordance with their values, beliefs, experiences and personalities. As a consequence, designing policy simulator systems to address such political phenomena require non-trivial field- and modelling work, because they involve a great range of behavioural responses and motivations.

This chapter will not focus on testing different *policy-schemas* to *improve* the state of the world according to the view of a policy-maker. Instead, this exercise aims at exploring how the impact of policy effects on stakeholders’ (value-driven) social acceptance can be modelled and simulated, and how this can lead to policy shifts and eventually different social outcomes. Given this purpose, the model in this chapter is an instantiation of the "meta-model" (to study a particular case study) and a possible computational implementation (Table 10.1). That said, the model in this chapter provides a generic ABM framework to cover these kind of social phenomena. Particular case studies would require specific data-gathering, modelling, and analysis to produce useful insights. Therefore, the model has some limitations. For instance, it does not consider many other relevant stakeholders involved in the domain, like water-intensive industries or businesses, and does not capture all possible actions *policy influencers* can perform. Because the model is a generic approach to these social phenomena, much simplification is necessary. Working on particular case studies allow a more accurate representation of the actual phenomena.

Table 10.1: Framework applied to the case study of policy shifts in the urban water domain

Domain (model and data)		Water & economy, Politics, Conservation practices (see Sec.10.4.7)
Stakeholders	**Policy-influencers**	Municipality, Water utility, Social movement, Political parties (see Appendix A)
	Policy-target	Households
Value model	**Policy-influencers**	Diverse profiles (*control*, *competence*, *efficiency*, *social justice*, *citizen involvement* and *water security*) (see Sec. 10.4.3).
	Policy-target	(Households): diverse profiles (*self-centred*, *universalist*, and *conservative*) (see Sec. 10.4.4).
Policy-schema		Diverse (see Appendix A and B)

In conclusion, this chapter provides a prototype to guide the modelling of those particular case studies related to social acceptance and policy shifts, focusing on the use of values. In particular, values play a role at two different levels: (i) as motivators for action of agents; and (ii) as social standards that drive the assessment of the state of affairs.

10.2 Introduction to the case study

According to the United Nations, by 2050, 68 % of the world's population is projected to live in urban regions [259]. Although urban water accounts for only 21 % of the water withdrawals in Europe nowadays [80], households are, in general, the main users of water of the public water supply, specially in those whose economy is based on services rather than manufacturing industry [79]. For instance, in Spain almost 70 % of the public water supply is for households [123].

This suggests that many citizens' direct well-being is linked to the governance and management of the public water supply. Households use water for diverse needs, such as hygiene, drinking, and sanitation. Noteworthy, the new challenges for cities (e.g. global warming, urbanisation, population growth, etc.) are producing phenomena that may impact citizens' well-being with regard to water (see [51, 266, 241]). In addition, some authors claim that economic inequality is rising, which may contribute to populism and polarisation of politics [184], which is expected to have consequences on environmental issues (see [98]).

Water public policies in urban regions must manage water services to address such collective challenges, while being in line with other services essential for citizens (e.g. housing, education, public health, etc.). However, addressing severe threats may rely on the enactment of mechanisms that can lead to social acceptance issues. For instance, economic instruments aiming at preserving a safe environment will not be sustainable if these also increase the social inequality.

Consequently, it is crucial to analyse the effects of public policies on social systems before the actual enactment of the policies. Public policies must preserve and improve the welfare of population in challenging situations that involve value dilemmas. Moreover, it is convenient to consider that multiple stakeholders will seek to produce policy shifts in order to ensure what they think is critical and imperative for citizens in such situations, which may produce unexpected outcomes.

Hence, the approach presented in this chapter is to build an agent-based model (ABM) to simulate and monitor the effects of public policies on social systems, considering the intervention of political actors to produce policy shifts, by using values.

10.3 Background

10.3.1 Water as an ethical and political space

Governance for water public services presents some particularities. For instance, there is practically no regular contact between households and water utilities outside of paying the water bill [45]. In fact, the average amount of time that households deal directly with their water utility is less than ten minutes per year, and the water service enters to their minds only when the tap water smells or tastes bad, or when the water bill arrives [199].

However, although such interaction between water utilities and households is limited, it has been acknowledged that the utility's conduct as a service provider is crucial for building public trust, confidence, and accountability [57]. Citizens want public services to be managed according to their ethical standards [168, 57]. Recently, the Committee on Standards in Public Life in United Kingdom alerted on the poor performance and "ethical failure" by private service providers, for which they recommended to guide management by the values, attitudes and behaviour the public expects in the delivery of public services [57, 58].

This suggests that the company's conduct, and not only the service, takes special relevance in citizens' minds. Moreover, since the water service is essential for citizens and is delivered as a natural monopoly (i.e. one entity manages the public water infrastructure) and where demand is, in general, inelastic, much of the social protest occurs in a political arena. In other words, households' beliefs about how the society and public services are not exhibited by making different choices in a market (i.e. consumer), but by participating and contesting such views in a political arena (i.e. citizens) [168].

That said, political behaviour of citizens has been acknowledged to be not as "rational" as many theories have proposed [3]: indeed, citizens are often poorly informed and use multiple heuristics and biases to make sense of the political world [152]. With this in mind, multiple stakeholders act within the political arena to dispute the view on how to govern and manage water services and to (presumably) execute citizens' demands. In addition, the water domain connects, politically, with other policy domains, meaning that those "ethical failures" in those domains produce effects in the political arena of the water-domain.

With this in mind, it is clear that much of the political action implies that there are agents that can recognise emerging macro-phenomena and, as a consequence, intentionally support or hinder the phenomena or the emerging process itself —which is known as *second-order emergence social phenomena* (*EP2*) [52, 210, 175] (see Chapter 5). Many political actors focus on assessing the state of affairs at the macro-level, and propose policy shifts or react to the occurring phenomena. On the contrary, most citizens perceive the water service only at the home level, and "consult" those political actors to make decisions concerning the state of affairs at the macro-level.

10.3.2 Water as a public service and values

Much literature that explores trust in public services approaches households as consumers, focusing on the client-company relationship. In this light, Sirdeshmukh *et al.* [238] distinguished two facets of consumer trust: (i) the *frontline employee behaviours*, that refers to the contact during a service encounter; and (ii) the *management practices and policies*, that refer to the guidelines that govern that exchange. According to the authors, three dimensions are important in these facets: (i) *operational competence* (i.e. perceiving that the company has the skills, ability and knowledge to perform the task effectively); (ii) *operational benevolence* (i.e. behaviours that exhibit that consumers' interests are ahead of company's interests); and (iii) *problem-solving orientation* (i.e. motivation to anticipate and solve problems that may arise during or after the service exchange). They also reported that trust and confidence are constructs that are gained or lost asymmetrically (e.g. negative experiences with frontline employees will typically decrease trust much more than positive experiences increase it) [238].

However, much confidence would come from the water utility's conduct at a social level, which implies, unavoidably, to represent the utility as a political actor that can affect the state of affairs. Given this issue, several frameworks propose multiple principles to deliver the water public service properly. For instance, the United Nations [260] recommend to strength the *availability*, *affordability*, *accessibility*, *safety* and *adequacy* of public services to ensure that human rights are realised and preserved.

In the model presented in this chapter different values will be explored. The reason is twofold: (i) it is presumed that other values away from a service-centric view are relevant for the stakeholders in the water domain (e.g. control, security, etc.); and (ii) it is acknowledged that many of these abstract values are difficult to be instantiated into concrete terms.

10.3.3 Households' values in the urban water domain

Values have been broadly studied in individuals' pro-environmental behaviour [67, 209]. Saurí [216] reported that water use in Spanish cities is declining because of (i) the adoption of water-saving technologies as well as behavioural changes caused by (ii) water awareness campaigns and (iii) rising water prices and taxes. Certainly, households' values and other cultural beliefs determine how relevant they consider these topics and how they react to persuading messages promoting water conservation. For instance, Owen [180] criticised that many pro-environmental campaigns relied on the fact that citizens would change their behaviour simply because they would have more knowledge and information, indicating that other social, personal, and political factors are involved in guiding their behaviour.

In the same vein, Jorgensen *et al.* [127] reviewed multiple models of household water use and concluded that much more knowledge about the diversity of factors involved in households' water use is needed. Consequently, factors related to use habits, motivations, institutional trust, and community trust are necessary for building better models. They identified trust to be key in water use: households will not save water if they believe that other households are not going to minimise their water use (i.e. *inter-personal trust*); and, likewise, they are less likely to save water if they do not trust the water authority (i.e. *institutional trust*).

Nevertheless, the purpose of this model is not to estimate accurately the water demand of households, but to explore the impact of different populations (i.e. holding different values) in the water-domain state of affairs. Accordingly, since households' values play a role in their political behaviour [152], different populations are expected to affect their response to public policies and, eventually, the social outcome. In fact, this is a research line that confronts the existence of an "average user" of water services and which considers that "soft" differences (e.g. lifestyles, personal views, values, norms, goals, motivations) in perspectives on water may be more useful than a classical segmentation of the population based on "hard" socio-demographic differences [41] —which could be applied to improve the citizen's involvement in water public services.

Building on these insights, the present model introduces the *satisfaction* of households as triggers for seeking political interventions (see [44]). It is given by the assessment of two components: (i) the water service, that is related to the fulfilment of their expectations in a domestic setting; and (ii) the social outcome, that is related to the social expectations of the public service. For the sake of simplification, the service assessment in (i) does not include all the facets identified by Sirdeshmukh *et al.* (see [238]). However, by introducing (ii), the model allows to include other elements in the social context into account, which are relevant to explore social acceptance issues.

10.3.4 Agent-based models for urban water

Agent-based models are increasingly used to approach urban water management. Mashhadi Ali *et al.* [158] developed a framework to explore the dynamics of water supply and water demand, modelling the interaction between the households and the water utility manager. Moss and Edmonds [170] studied the effects of social influence on domestic water demand. Galán *et al.* [88] used an ABM as a complementary tool to examine the causal relationships present in the water demand in metropolitan areas (e.g. opinion diffusion, technologies adoption, urban dynamics, etc.). Schwarz and Ernst [228] modelled the diffusion of water-saving technologies by clustering the population into five profiles in accordance with their lifestyles and values. Noticeably, these models focus, generally, on one of the main topics of the urban water management: forecasting the domestic water demand.

In contrast, the purpose of this model is to simulate and monitor the effects of public policies on social systems, considering the participation of political actors and using values, and identify the main elements that are necessary for modelling such cases.

One of the required affordances for agents that represent political actors is to perceive, assess, and react to second-order phenomena. These phenomena refers to the idea to recognise emerging macro-phenomena and, as a consequence, intentionally support or hinder the phenomena or the emerging process itself [52, 210, 175]. This affordance is useful to describe the capabilities of political actors, who, in general terms, focus on outcomes at the social level, rather than effects on isolated individuals. In multi-agent systems, this notion has been used to explore the emergence of reputation, through the spread of rumours in a social group [175]. In agent-based simulation, it underlies models that explore the creation of entities to prevent or support social trends (for instance, Vanhée *et al.* [263] presented a model of a tribal society in which violence triggered the creation of social-control organisations in order to prevent it, although they emerged as a consequence of locally-suffered attacks and not due to the observation of macro-phenomena).

With regard of addressing phenomena in political arenas, this is an issue that has been not much explored in agent-based models (e.g. electoral outcomes and voting behaviour and social effects, etc.). Instead, these have usually been approached with game theory [155]. Typically, games model a few agents that choose optimally between possible strategies, and are used to derive in which conditions the equilibrium exists. However, agent-based models admit a richer representation of the real-world (e.g. social networks, geography, richer heterogeneity, etc.).

In this line, *social acceptance* is a topic that has been identified as relevant for public policies. De Wildt *et al.* [269] pointed out that *social acceptance* is key for policy-making, as it may jeopardise the achievement of the goals of the policy. They studied social acceptance of smart electricity grid deployment focusing on values, indicating that unsatisfied expectations concerning values may eventually result in social acceptance issues (although it is clear that the underlying causality of social acceptance issues is more complex). Nevertheless, no agent-based models have approached *social acceptance* using a value perspective, to the knowledge of the author of this dissertation. The model presented in this chapter aims at addressing such gap, by using the theoretical notions described in Chapter 5. Besides modelling the stakeholders' satisfaction concerning values, this model takes one step further by modelling their potential reactions (in a hypothetical situation).

10.4 Model

The purpose of the model is to test policies and then observe their effects on the socio-economic environment taking the policy shifts produced by *policy-influencers* into account. Hence, the model affords to explore the social acceptance of policies through the stakeholders' degree of satisfaction of those expectations concerning values.

The model represents a urban population constituted of households and policy influencers. The model simulates one decade of activity through discrete time-steps of one month. Each month households (i.e. *policy-targets*) demand water, receive the water bill, may adopt conservation practices, assess their satisfaction, and may support political demands. Likewise, *policy-influencers* evaluate the state of the world and may advocate for political interventions.

10.4.1 Value model

The model has been built with the following (social) values in mind —that is, the model has components that are presumed to reflect these values, following the methodology explained in the previous chapter. Furthermore, each agent has its own *value profile* (see Secs 10.4.3 and 10.4.4 and Appendix A).

V1. **Control.** The model considers *control* as the share of public participation in the water utility, since the public administration is the holder and ultimate responsible of the water service. It is a value that multiple stakeholders in the policy domain consider essential for the urban water services.
V2. **Competence.** The model affords to observe and manipulate variables related to the competence in managing the water service from an economic perspective (i.e. financial sustainability of the water service through water fees).
V3. **Efficiency.** The model affords to observe and manipulate variables related to the (responsible) use that is made of water to met the citizens' water needs (i.e. households' water use and water losses).
V4. **Social justice.** The model affords to observe and manipulate variables related to the vulnerability of the households with regard to the water supply (i.e. affordability of the water fees, households' debts, service shutoffs, and households that benefit from financial support).
V5. **Citizen involvement.** The model affords to reflect the participation of households when supporting policy influencers' demands.
V6. **Water security.** The model affords to observe and manipulate variables related to the water security (i.e. long-term sustainability) in an urban context (i.e. households' water use).

10.4.2 Agent roles

Two main roles are considered for the artificial agents: (a) *policy-targets* and (b) *policy-influencers* (see Chapter 5). The *policy-influencer* role is divided into more specific sub-roles (Table 10.2) (see Appendix A).

Table 10.2: Agent sub-roles considered in the model for simulating policy shifts in the urban water domain

Policy-targets	**Policy-influencers**
— Households	— Municipality — Water utility — Social movement — Political party "NL" — Political party "SL" — Political party "S"

10.4.3 Policy-influencers

Policy-influencers are characterised by (i) a *value profile*; (ii) their *political satisfaction*; (iii) a *set of actions*; and (iv) a *set of political demands*.

Value profile

Values represent what *policy-influencers* consider relevant in the world, thus driving the assessment of the state of the world and determining their political demands.

A vector $\vec{V} = (V_1, V_2, ..., V_n)$ is used to represent the set of values held by a *policy-influencer*, whose components V_i correspond to value items and whose score may be 1 or *true* (i.e. the value item is held) or, 0 or *false* (i.e. the value item is not held) (Table 10.3). It is convenient to remark that "not holding a value" does not mean that the *policy-influencer* considers that the value is irrelevant or a disvalue. Rather, it means that a particular value is not an actual *motivator* of their demands/actions in that particular context, and thus unsatisfied expectations with regard to this value do not *trigger* political action. For instance, a *policy-influencer* who does not hold the value of *environmental protection* does not claim it to be a disvalue, but rather, it represents that the agent does not evaluate the world-state with regard to that value and, as a consequence, does not perform any action/demand that pursues to improve the world-state in connection with *environmental protection*. However, even in this case, the *policy-influencer* can frame their political demands to appeal to *environmental protection* (although they were not motivated by that particular value) (see Sec. 10.4.8).

Held values are ordered by relative importance (that is, there is an explicit value hierarchy). This is represented with a vector of value weights $\vec{\beta} = (\beta_1, \beta_2, ..., \beta_n)$. It is assumed that, when a value j is not held, its weight is null (i.e. $V_j = 0 \rightarrow \beta_j = 0$).

When the value j is held, the *policy-influencer* evaluates the world-state W using a set of value functions $\boldsymbol{E}^{(j)}$. This results in an evaluation score e_j with regard to that particular value j (see Sec. 10.4.7). Afterwards, the *policy-influencer* aggregates the set of value scores $\boldsymbol{e}$ into one overall evaluation e_T using their respective weights $\vec{\beta}$, which is used to determine its *political satisfaction*. The scores (and, by extension, also *political satisfaction*) range from 0 (totally displeased) to 1 (completely satisfied).

Table 10.3: Generic value profile of a *policy-influencer* in the model

Value item	Is it held?	Weight	Evaluation score
Value 1	*true* or *false*	β_1	$\boldsymbol{E}^{(1)}(W) = e^{(1)} \in [0,1]$
{...}	{...}	{...}	{...}
Value n	*true* or *false*	β_n	$\boldsymbol{E}^{(n)}(W) = e^{(n)} \in [0,1]$

Actions

Policy-influencers have a set of “generic actions”:

- — Perceive (a fragment of) the world-state (see Sec. 10.4.7).
- — Evaluate (a fragment of) the world-state (see Sec. 10.4.7).
- — Push for/Withdraw a *political demand* (see Sec. 10.4.8).
- — Inform and influence households (See Sec. 10.4.7 and Table 10.10).
- — Interact with other *policy-influencers* (see Sec. 10.4.2).

These generic actions have to be specified for the multiple sub-roles defined (for instance, the *water utility* does not perceive and evaluate the same aspects of the state of the world as the *social movement*). In the same vein, the specific preconditions of advocating for a *political demand* depends, clearly, on the actual demand. As a consequence, the specific *set of actions* depends on the actual *policy-influencer*.

Political demands

Policy-influencers may advocate for *political demands* when their evaluation of the world-state is opportune, in order to change it towards their desired outcomes, which are defined by their values.

With this in mind, each *political demand* has *preconditions* that, if met, it is raised in the social space, meaning that all the agents are aware of that demand. For instance, households may support it, “forcing” the *municipality* to enact instruments in correspondence with that matter. These set of preconditions ***C*** are based on the set of evaluations of the world-state ***e*** with regard to particular values, which represent the motives needed to act (in more precise terms, when the evaluation of the world ***e*** of the *policy-influencer* satisfies the preconditions of the *political demand*, it can be raised). For instance, a *policy-influencer* may have “*increase water fees*” among its *set of political demands*, which is demanded when the evaluation with regard to *competence* is opportune (e.g. the financial sustainability of the public service by means of water fees is not satisfactory).

Two types of *political demands* are considered:

(i) **Advocated demands.** These *political demands* are raised in the social space to gather support of other agents. They are passed to the *municipality*, advocating for a policy shift that takes the demand into account. It is assumed that they are enacted by the *municipality* when they gather enough social support from households.

(ii) **Enacted demands.** These *political demands* can be directly enacted in the social space without the intervention of the *municipality*, meaning that *policy-influencers* have the capacity to enact them (with no more *preconditions* than the ones related to the assessment of the world-state).

(iii) **Advocated + enactable demands.** These *political demands* can be directly enacted without the intervention of the *municipality*, but some additional preconditions may be required (e.g. some level of *social support*).

In this model it is assumed that raising *political demands* has no "cost" (e.g. economic, political, etc.) beyond an opportunity cost (that is, the amount of time invested in advocating for a particular demand is not used for advocating for another one).

Sub-roles

The *policy-influencers* included in the model are listed in the Table 10.4, as well as the values they hold. See Appendix A for a complete view of the *policy-influencers* (values, actions, and demands).

Table 10.4: *Policy-influencers* in the model and their held-values

Policy-influencer	**V_1= Control**	**V_2= Competence**	**V_3= Efficiency**	**V_4= Social justice**	**V_5= Citizen involvement**	**V_6= Water security**
Municipality	1	1	0	0	1	1
Water utility	0	1	1	0	0	0
Social movement	1	0	0	1	1	0
Political party "NL"	0	1	0	0	0	0
Political party "SL"	0	1	0	1	0	0
Political party "S"	1	0	0	1	0	0

10.4.4 Households

Households are characterised by (i) *value profile*; (ii) *number of members*; (iii) *income*; (iv) *debt-related variables* (i.e. *total debt* and *overdue-counter*); (v) *water-related variables*; (vi) *conservation practices*; (vii) *service satisfaction*; and (viii) *political satisfaction*.

The *value profile* defines the behaviour of *households*, namely, by means of specific decision-making models, value aggregation frameworks, and interaction with *policy-influencers* (i.e. from perceived trustworthiness to "goodness" of their political proposals) (see Sec. 10.4.7 and 10.4.8). It also determines their held values $\vec{V}$.

The *water-related variables* include: *water demand* (i.e. the volume of water that is requested); *water use* (i.e. the volume of water that is actually used); *revenue water* (i.e. the volume of water that is billed); and *water bill* (i.e. the amount of money to be paid, in accordance with the *revenue water*). These variables depend notably on the households' *number of members* and *conservation practices*. Moreover, they may affect the *service satisfaction* (e.g. high *water bills* may decrease the *service satisfaction*) and households' *debt* (e.g. high *water bills* may make low-income households unable to pay, thus increasing their *debt*).

Elements (ii) and (iii) are based on real-data of one (generic) urban district (in particular, Barcelona's district *Esquerra del eixample* [122, 120]), (iv–ix) evolve as results of the simulation, and (i) is an input set by the user.

Value profiles

There are three different *value profiles* that households may have:

- ***Self-centred households***, whose dominant values are related to *self-enhancement* and *individualism*. These households think that the service should cost little money and should respect households' autonomy (which they consider it is well represented by low water fees and no water use restrictions). Their (social) held-value is only *competence*, as they value cost reductions rather than additional services or initiatives.
- ***Conservative households***, whose dominant values are related to *tradition*, *security* and *conformity*. These households think that the public service must be guaranteed in any case, and they understand that this means that some situations may require restrictions must be applied to ensure their financial and ecological sustainability. Their (social) held-values are *competence*, *efficiency*, and *water security*.
- ***Universalist households***, whose dominant values are related to *universalism*. These households think that the service must promote social justice (i.e. it must take vulnerable households into account) and must be completely accountable and open to citizens' opinions. They need to know that the water they use is as sustainably and ethically sourced as possible. Their (social) held-values are *control*, *social justice* and *citizen involvement*.

10.4.5 Initialisation

Households start with an input *value profile* and the least efficient *conservation practices* in all the domestic uses. The *number of members* is randomly assigned, based on actual statistical-data of one urban district (in order to use real information, available data of the district of *Esquerra del Eixample* has been taken [122]). The *income* is given by the annual income of households depending on their *number of members* (available in [120] for the same district); as this information is given by income ranges, the actual income is randomly generated following a uniform distribution. Notice that the *number of members* and *income* are randomly set regardless of the *value profile* of households. All households start with no *debt*.

All households are linked to all the *policy-influencers*, and therefore they can be informed and influenced by these. The agent population is constituted by 100 households and the (six) *policy-influencers* mentioned in Table 10.4.

The service cost is set to 2.20 eur/m^3 (see [5]). For the sake of simplicity, the total service cost is assumed to increase linearly with the total volume of supplied water. The tariff-multiplier factor starts at $\gamma = 1$ (see Sec. 10.4.7).

10.4.6 Process Overview

The following procedures are executed sequentially each time-step (one month) (see Secs. 10.4.7 and 10.4.8 for further details):

P1. **Water supply cycle:**

  P1.1. **Households demand water**, according to their domestic needs and conservation practices.
  P1.2. **Households receive water**, which is registered by a water-meter, determining the water bill to be paid. The *water utility* supplies water to the households, as long as their supply has not been shut off.

  - P1.3. **Households pay the bill**, as long as it is affordable, according to the households' income. When the water bill is too high, they might not pay, which increases their debt. In this case, they will be requested to pay part of their debt in the following months (which may make following water bills more difficult to be paid). If households are able to pay, their debt decreases. When they have accumulated three months without paying the water bill, the water utility may shut their water supply off (this period of time is set for testing purposes, given the absence of reliable data).
  - P1.4. **Households adopt (or abandon) conservation practices**. If the bill is high, they may adopt a practice in some household setting (e.g. shower or laundry). On the contrary, if the bill is low, they may abandon a conservation practice. Environmentalist households may also adopt practices when their water use is too high.
  - P1.5. **Households apply for social aid**, as long as it is enabled. Thus, households must ask for financial assistance, if they want to, and must prove that they are potential beneficiaries (namely, their income is lower than 13,000 eur/year, according to [82]).
  - P1.6. **Households evaluate the service**, comparing the water volume they received with the volume they requested, and assessing the cost of the service.

- P2. **Social and political:**
  - P2.1. **Policy-influencers evaluate the state of the world** according to their values, that is, focusing on some particular variables of the world.
  - P2.2. **Households evaluate the political state of the world**, retrieving information and opinions from the *policy-influencers*.
  - P2.3. **Households global assessment**, for which they aggregate the service and the political evaluations.
  - P2.4. **Policy-influencers advocate for political demands**, according to their evaluations of the state of the world.
  - P2.5. **Households support political demands** among those that have been raised in the social space.
  - P2.6. **Policy-influencers enact demands**, as long as they are afforded to do so (e.g. they gathered enough social support), and the conditions of the state of the world are opportune.
  - P2.7. **Policy-influencers request to the *municipality* to enact their demands**, as long as these have gathered enough social support.

10.4.7 Submodels

Water demand and conservation practices

Conservation practices represent the behaviour of households with regard of the multiple domestic water uses. They are expressed by a numerical score that can be 0 (least efficient practice in that context), 1 (moderately efficient practice), or 2 (most efficient practice). *Conservation practices*, as well as the *number of members*, determine the *water demand* of households each time step. Multiple sources have been used to define these practices (e.g. [69, 78]).

In general, *conservation practices* can be adopted and abandoned. However, it is assumed that those stages that represent technology adoptions (e.g. adoption of water-saving shower heads or new washing machines) cannot be abandoned.

Domestic water uses are:

- **General hygiene:** water that members make use of to do hygiene-related activities (e.g. washing their hands, brushing their teeth, etc.).
- **Toilet:** water that members make use of to (mainly) discharge organic waste of physiologic origin.
- **Dishes:** water that households make use of to washing and rinsing dishes and kitchenware. It is assumed that a household do the dishes for all its members.
- **Laundry:** water that households make use of to wash clothes. It is assumed that a household does the laundry for all its members.
- **Shower:** water that members make use of to wash their bodies. It is assumed that every member in the household has a shower once a day.
- **Direct consumption:** water that households make use of to cook or to drink. It is assumed that the daily demand for direct consumption is 15 litres per household (L/hh) (no conservation practices).
- **Household maintenance:** water that households make use of to clean the house and to irrigate plants —but no gardens. It is assumed that the weekly demand for household maintenance is 50 L/hh (no conservation practices).

The *conservation practices* for each domestic use are listed in Table 10.5.

Water billing

This submodel is used to calculate the water bill of households of one month, following the tariff scheme of *Aigües de Barcelona* [7].

In general terms, the domestic water use is charged by means of two components: a fixed service fee, that depends on the housing nominal flow (e.g. 0.40 m^3/h); and a variable amount, that depends on the volume of used water. Additionally, the water bill may include tributes and taxes.

In this model, it is assumed that all households have the same type of tariffs (nominal flow of 0.40 m^3/h), meaning that the service-fee is 7.54 eur/month for every household.

The variable concept is based on a block tariff, which applies a progressive fee according to the volume range of water used. The volume ranges and water prices are retrieved from *Aigües de Barcelona* (Table 10.6), which is the metropolitan water utility (see [7]). The volume ranges are set for a household of three members; if the household is larger, it is possible to widen their ranges (for instance, 2 m^3/month per additional member for the first block).[1]

Moreover, the Catalan Water Agency applies a regional environmental tribute, which is also calculated with the blocks scheme (Table 10.7) (see [4]). Noteworthy, it charges at least 6 m^3 although the actual water use is lower, and its ranges can be widened if more than three people live in the household (3 m^3/month per additional member). Finally, a value-added tax of 10 % is added.

If social aid is enabled, those households that have been identified as vulnerable may benefit from a lower water price and a lower water tribute. This bonus is applied as long as the household's water use does not exceed the first volume range (see [7]). Accordingly, the first block is updated as follows: the water price is the 75 % of the regular price, and the tribute is reduced to 0.24 eur/m^3.

---

[1] In 2020, the system for readjusting the blocks when more than three people live in the household has changed. See `https://www.aiguesdebarcelona.cat/ca/el-teu-servei-daigua/bonificacions-i-fons-de-solidaritat/` (in Catalan)

Table 10.5: Conservation practices of households depending on the domestic water use

Use	**Least efficient (0)**	**Moderately efficient (1)**	**Most efficient (2)**
General hygiene	Members use faucets during 3 minutes a day with a flow of 10 L/min, keeping them on unnecessarily. The daily demand is 30 L/p.	The household adopts faucet aerators. Members use faucets during 3 minutes a day with a flow of 5 L/min, but they still keep them on unnecessarily. The daily demand is 15 L/p.	Members use efficiently faucets, during 1 minute a day with a flow of 5 L/min. The daily demand is 5 L/p.
Toilet	Members not only use the toilet for physiologic reasons but also to flush other solid waste. A flush requires 10 L and every dweller flushes the toilet six times per day. The daily demand is 60 L/p.	Members use the toilet only for physiologic reasons. A flush requires 10 L and every dweller flushes the toilet four times per day. The daily demand is 40 L/p.	Members use the toilet only for physiologic reasons and they use properly the water-saving flushes. It is considered that a flush requires on average 6 L and every dweller flushes the toilet four times per day. The daily demand is 24 L/p.
Dishes	Households do the dishes by hand twice a day, requiring 75 L per wash. The daily demand is 150 L/hh.	Households do the dishes by hand twice a day, requiring 30 L per wash. The daily demand is 60 L/hh.	Households use a dishwasher once a day that requires 15 L per wash. The weekly demand is 15 L/hh.
Laundry	Households use an old washing machine that requires 130 L/w per wash twice a week. The weekly demand is 260 L/hh.	Households use a new washing machine that requires 60 L/w per wash twice a week. The weekly demand is 120 L/hh.	Households use a new washing machine that requires 60 L/w per wash once a week. The weekly demand is 60 L/hh.
Shower	Every dweller has a shower of 6 minutes and a flow of 15 L/min. The daily demand is 90 L/p·d.	The household adopts a water-saving shower head. Every dweller has a shower of 6 minutes and a flow of 10 L/min. The daily demand is 60 L/p.	Every dweller has a shower of 4 minutes and a flow of 10 L/min. The daily demand is 40 L/p·d.
Direct consumption	Members use faucets to cook and drink. The daily demand is 15 L/hh.		
Household maintenance	Members use faucets for cleaning and irrigating indoor plants. The weekly demand is 50 L/hh.		

Table 10.6: Water use pricing (Source: [7])

Block	**Range (m^3/month)**	**Price (eur/m^3)**
1	[0, 6]	0.6087
2	(6, 9]	1.2175
3	(9, 15]	1.8262
4	(15, 18]	2.4349
5	>18	3.0436

To change water prices, the model includes a factor γ that multiply water fees and tributes, but keeping constant the volume ranges. Thus, a value for this fee-factor γ lower than 1 decreases the water

Table 10.7: Water regional tribute (Source: [4])

Block	**Range (m^3/month)**	**Price (eur/m^3)**
1	[0, 9]	0.4936
2	(9, 15]	1.1370
3	(15, 18]	2.8425
4	>18	4.5480

fees, while a value greater than 1 increases them.

The input of this submodel is the revenue water of one month (i.e the water volume, in cubic meters, that has been billed) and the output is the payment due. The payment can be considered as a relative bill if compared to the household monthly income (%).

Households pay the water bill

Households pay the water bills as long as it does not exceed the 10 % of their monthly income (given the lack of more detailed information). In case it does, it is possible that households cannot afford to pay for the water service. In absence of fieldwork, it is assumed that households will pay 50 % of these occasions.

When household cannot pay, its debt with the water utility grows (namely, the water bill is added to any previous debt) and its service-assessment drops to 0 (see Sec. 10.4.7). Besides, its *overdue-counter* increases by one: if the interruption of the water supply is allowed, a household whose overdue-counter is three or higher will have their supply shut off.

Having a debt with their water utility makes more difficult for household to pay future water bills. It is assumed that households have to cover a part of the debt incurred in the past. The minimum amount is the payment associated to 3 m^3 per household member (or, alternatively, the total debt if it is lower). This minimum payment is added to the next water bill, which can lead to exceeding the 10 % threshold aforementioned. If the household is nevertheless able to pay, its debt decreases by the minimum payment and its overdue-counter is reduced by one.

Those households whose supply was shut off can restore it as long as they are able to pay the minimum payment (3 m^3 per household member) —provided it does not exceed the 10 % threshold. As before, their debt decreases accordingly and their overdue-counter is reduced by one.

Some *policy-influencers* may mobilise the households to declare a massive refusal to pay the water bills, as an expression of social protest. In this case, those households that support this form of social protest do not pay the water bills. It is assumed that, since it is a massive protest, their debt does not increase and their supply cannot be shut off (for instance, households appeal in mass against the water bill, through some entity or collective, and they collapse the legal department of the water utility). *Conservative* households cannot support this social protest.

Households apply for social aid and adopt conservation practices

Both applying for social aid and adopting conservation practices depend, in general, on the value profile of households (Table 10.8). When households have a debt with the water utility, or when the supplied water is below their water demand (namely, 80 %, for the sake of testing the model), they will adapt to these conditions by trying to adopt one conservation practice —in case of debt, they are more likely to adopt conservation habits rather than buying household appliances.

Table 10.8: Behaviours when applying for social aid and adopting conservation practices according to the value profile of the household.Given the absence of accurate data, thresholds have been assumed for test purposes.

Value profile	Conservation practices	Application for social aid
Self-centred households	Adoption only due to high water bills (namely, when it is greater than 6.5 % of their monthly income). They will abandon such practices if the bill is low enough (when it is lower than 1.5 % of their monthly income). The likelihood that they buy a new household appliance (e.g. dishwasher or new washing machine) is greater than adopting a pro-environmental habit (e.g. shorter showers or avoid unnecessary toilet flushes).	In the case that the social aid is enabled, they will apply only after their water supply has been shut off.
Conservative households	Adoption only due to high water bills (when it is greater than 5 % of their monthly income). They will abandon such practices if their bill is low enough (when it is lower than 0.5 % of their monthly income). They are more prone to adopt a pro-environmental habit than to buy household appliances.	In the case that the social aid is enabled, these households will apply for social assistance when they have low satisfaction with regard to the water bill, or alternatively they have been two months without paying for the service.
Universalist households	Adoption because of economic (when the bill is greater than 5 % of their monthly income) and environmentalist reasons (when their water use is greater than 150 litres per person and day). They may drop conservation practices if their bill is low enough (lower than 1.5 % their monthly income). The likelihood that they buy a new household appliance is the same as adopting a pro-environmental habit.	In the case that the social aid is enabled, these households will apply for social assistance when they have low satisfaction with regard to the water bill, or alternatively they have been two months without paying for the service.

The adoption of a conservation practice is implemented as follows: when the adoption condition is met (for instance, the water bill is high), the household will choose one domestic use at random, whose likelihoods depend on the household's value profile. If their behaviour can be improved (that is, their *conservation practice* is 0 or 1), their *conservation practice* for that domestic use will increase by one. The likelihood to select one domestic use that involves the adoption of household appliances may be greater for certain value profiles —representing that some households are more prone to buy efficient technology rather than adopt habits. Notice that this may lead to "unsuccessful attempts", since they may choose uses in which they are already efficient.

Households' satisfaction

Households' satisfaction is constituted by two components:

— *Service assessment* based on (i) volume supplied to the household; and (ii) water bill. These are local variables that households perceive and can evaluate directly. Other aspects that could be taken into account are water quality (mainly, organoleptic properties of water, such as taste or colour), water pressure, supply interruptions (e.g. due to maintenance works), billing accuracy (the billed volume corresponds to the volume that has been registered by the water meter), or client-company relationship (e.g. behaviour of frontline employees).

(i) *Supplied-volume assessment*, that compares the volume received with the volume they requested (Fig. 10.1);

(ii) *Water bill assessment*, that focus on the cost of the service with regard of the household income (Fig. 10.2).

— *Political assessment*, that focuses on the state of the world according to the values they hold. These are macro-level variables that households perceive and evaluate indirectly, through *policy-influencers* (see Chapter 5). Messages from *policy-influencers* contain their aggregated political evaluation, being used by households to form their own view, which is derived from the influencers' *political-satisfaction*. In particular, this component is the result of a weighted mean of the policy-influencers' *political-satisfaction* using the *value similarity* λ as weight, in order to generate an average opinion (see Sec. 10.4.8).

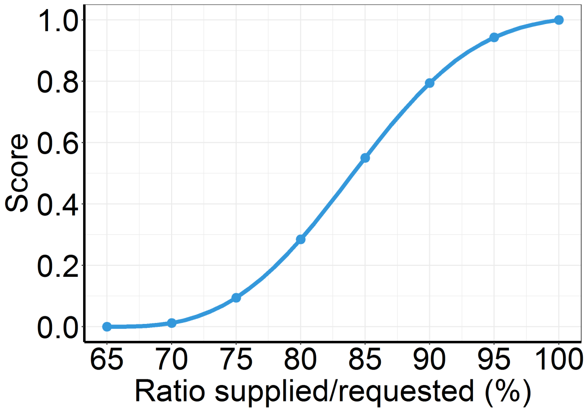


Figure 10.1: Households' value function for *supplied-volume assessment*. As shown, the score decreases the more supply interruptions or restrictions the household experiences (in general terms, the household can use as much water as wanted). The score is acceptable (0.7–1.0) when the ratio of water supplied:water requested is more than approximately 90 %. After that ratio, the score decreases rapidly. Ratios below 75 % are unacceptable, resulting in low scores (0.0–0.2).

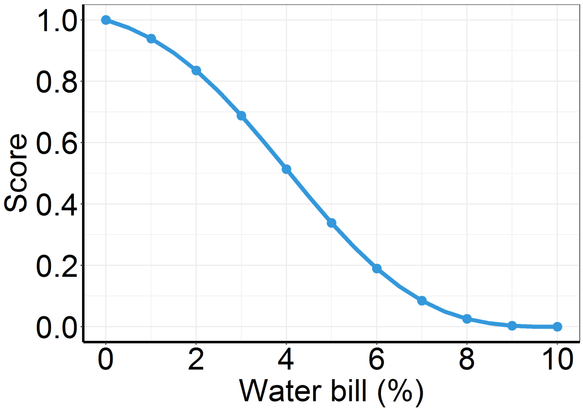


Figure 10.2: Households' value function for *water bill assessment*. As shown, the score decreases as the more expensive is the water bill. The score is acceptable (0.7–1.0) when the bill is around 0 and 3 % the households' income (UN considers the service *affordable* when the cost does not exceed 3 % of household income [257]). After that, the score decreases rapidly. When the service cost doubles the *affordable* threshold (i.e. 6 %) the score is very low (0.0–0.2).

The *supplied-volume assessment* and *water bill assessment* in *service assessment* are aggregated using a geometric mean. The *global satisfaction* (i.e. *service assessment* and *political assessment*) is also aggregated using a geometric mean. The reason to use the geometric mean is that, although components are equally weighted, low scores of one component cannot be compensated by high scores of another one.

Value functions for political assessment

Value functions are used as *evaluation mechanisms* to determine the *political assessment* of *policy-influencers* with regard of their values, as presented in Chapter 4. Each time-step the *policy-influencers* evaluate the world-state. If they hold a value, they have, at least, one corresponding value function for assessing the state of the world with respect to that particular value. Consequently, they generate a set of evaluations ***e***, that are aggregated using a weighted arithmetic mean (in order to have more stable assessments).

Households' value similarity with policy-influencers

This submodel is used to compute the "sympathy" of households for *policy-influencers* by means of their value similarity. The *value similarity* λ ranges from 0 (i.e. no similarity at all) to 1 (i.e. perfect similarity).

The *value similarity* is computed as follows (Alg. 1): each value i is compared; if the *household* holds the value, and also does the *policy-influencer*, their similarity increases by one. The same happens if the *policy-influencer* holds the value, and also does the *household*. At the end, the *value similarity* is normalised, taking only those values that are held either by the *household* or the *policy-influencer* into account.

Algorithm 1: Compute the *value similarity* between a household and a policy influencer

```
Input: Policy-influencer's held-values $\vec{V}^{(PI)}$, Household's held-values $\vec{V}^{(H)}$
Output: Value similarity $\lambda$
foreach pair $\{V_i^{(PI)}, V_i^{(H)}\}$ with $(V_i^{(PI)} = 1$ or $V_i^{(H)} = 1)$ for $i = 1..n$ do
    $m \leftarrow m + 1$
    if $V_i^{(PI)} = V_i^{(H)}$ then
        $k \leftarrow k + 1$
return $\lambda \leftarrow k/m$
```

Value similarity λ is used to define the weight of each *policy-influencers' political satisfaction* when determining the households' *political satisfaction* (see Sec. 10.4.7). Accordingly, the *value similarity* between the different *value profiles* of households and the multiple *policy-influencers* is shown in Table 10.9.

10.4.8 Policy-influencers advocation

Policy-influencers may advocate for *political demands* when their evaluation of the world-state is opportune, in order to change it towards their desired outcomes, which are defined by their values.

With this in mind, each *political demand* has *preconditions* that, if met, is raised in the social space, meaning that all the agents are aware of that demand. For instance, households may support it, "forcing" the *municipality* to enact instruments in correspondence with that matter. These set of preconditions ***C*** are based on the set of evaluations of the world-state ***E*** with regard to particular values, which represent the motives needed to act (in more precise terms, when the evaluation of the world ***E*** of the *policy-influencer* satisfies the preconditions of the *political demand*, it can be raised). For instance, a *political party* may have "*increase water fees*" among its *set of political demands*, and it is demanded

Table 10.9: *Value similarity* between *households* and *policy-influencers*

Household Profile / **Policy-influencer**	**Self-centred household**	**Conservative household**	**Universalist household**
Municipality	0.25	0.40	0.40
Water utility	0.50	0.67	0.00
Social movement	0.00	0.00	1.00
Political party "NL"	1.00	0.33	0.00
Political party "SL"	0.50	0.25	0.25
Political party "S"	0.00	0.00	0.67

when the evaluation with regard to *competence* is opportune (i.e. the financial sustainability of the public service by means of water fees is not satisfactory). The fact that all agents are aware of any *active political demands* is an assumption of the model, but other models could explore other social models (e.g. *policy-targets* can be disconnected from some *policy-influencers*, making them unable to perceive and support their *political demands*).

The set of political demands are implemented as follows (Table 10.10):

- **Political demand**, which is the identification of the particular action/demand.
- **Preconditions**, which are value-based preconditions (i.e. from the evaluation of the world-state) for performing the action or advocating for the demand.
- **Additional preconditions**, since some demands/actions may have extra prerequisites to be raised or enacted (e.g. social support). It is convenient to distinguish between additional prerequisites to raise the demand in the social space and additional prerequisites to enact the demand in the social space (see below).
- **Framing**, which represents the "discourse" with which is presented the demand in the social space to gather the support of households.
- **Type**, which the way by which the political demand can be enacted. Three types of *political demands* are considered:
  - (i) **Advocated demands.** These *political demands* are raised in the social space to gather support of other agents. To be enacted, they must be validated by the *municipality*, assessing their *social support*.
  - (ii) **Enacted demands.** These *political demands* can be directly enacted without the intervention of the *municipality*, meaning that *policy-influencers* have the capacity to enact them.
  - (iii) **Advocated + enactable demands.** These *political demands* can be directly enacted without the intervention of the *municipality*, but some additional preconditions may be required (e.g. some level of *social support*).

  Notice that all *advocacy demands* have *framing*, but only some *enactable demands* have so (since some of these will be directly enacted without "asking" for social support).

Policy-influencers check the world-state to define their *political demands* according to Alg. 2. Its output is a subset of their *political demands*, namely, those *political demands* that are appropriate for that particular world-state. In this model, it is considered that a *policy-influencer* can only advocate for one

Table 10.10: Generic structure of the set of political demands of a *policy-influencer*

Political Demand	**Preconditions**	**Additional**	**Framing**	**Type**
Demand d_1	$C_{d_1} = \{C_1^{d_1}, ..., C_n^{d_1}\}$	Add. preconditions R_{d_1}	$\vec{F_{d1}} = (F_1^{d_1}, ..., F_n^{d_1})$	Type_{d_1}
{...}	{...}	{...}	{...}	{...}
Demand d_k	$C_{d_k} = \{C_1^{d_k}, ..., C_n^{d_k}\}$	Add. preconditions R_{d_k}	$\vec{F_{d_k}} = (F_1^{d_k}, ..., F_n^{d_k})$	Type_{d_k}

demand at the same time, which requires to "select" among the subset of potential *political demands*. Several options could be used for this purpose (e.g. random selection, ranking, urgency, etc.). In this model, *political demands* are introduced by ranking (taking the value hierarchies and he effects of the demands into consideration), as presented, for instance, in Table 10.13; thus, the *policy-influencer* will always choose the demand that is "closer to the top" (i.e. reflecting the priority and urgency of the several options to be taken) (see Appendix A).

```
Algorithm 2: Compute the preselection of potential political demands
Input: Set of political demands D (preconditions C; additional advocacy-prerequisites R),
       Set of policy influencer's evaluations e
Output: Subset of potential political demands P
set P ← Ø
foreach demand d in D do
    if (preconditions C_d are satisfied by evaluations e) and (prerequisites R_d are met) then
        Put d into P
```

Once a particular *demand* has been raised, it becomes active in the social space (which is implemented as shown in Table 10.11). *Policy-influencers* cannot withdraw it until six time-steps have passed, which is when they evaluate the world-state again and consider opportune *political demands* again. It is assumed that the *political demands* have no political cost (in terms of consequences on reputation, investment of economic resources, etc.) beyond an opportunity cost (that is, the amount of time invested in advocating for a particular *demand* is not used for advocating for another one).

Table 10.11: Generic structure of the set of active demands in the social space

Active demands	**Social support**
Demand d_i	$O_{d_i} \in [0, 100]$
{....}	{....}
Demand d_j	$O_{d_j} \in [0, 100]$

Some of the *demands* can be enacted by the *policy-influencers* (i.e. *enactable demands*), which produce and effect on the social space. In this case, *policy-influencers* check each time-step whether their additional enactment-prerequisites are met (e.g. social support). If so, the demand is enacted an is withdrawn from the social space (i.e. Table 10.11). Notice that multiple *policy-influencers* may advocate for the same *political demand*; in this case, they gather social support using their own *framing* —which may help to reach different audiences.

Households' support political demands

This submodel is used to determine the social support given by households for those *political demands* that have been raised by *policy-influencers*, based on their affinity.

Demands are framed by *policy-influencers* to fit with how households view the world. In blunt terms, framing consists in how a certain political intervention is presented in the social space (e.g. concepts, metaphors, values, discourse) to the other agents in order to receive support. For instance, a proposal to "*pedestrianize a street*" may be framed as a "*measure to promote sustainable mobility*" (thus appealing to environmentalist values) or, alternatively, as a "*measure to promote the economic activity of the commercial street*" (thus appealing to wealth-related values). Nonetheless, the material intervention is the same regardless of its framing: in the previous example, a norm will be enacted to ban motor vehicles from circulating in that street.

Accordingly, *advocacy demands* have a *framing* $\vec{F} = (F_1, F_2, ..., F_n)$, whose components F_i correspond to the values aforementioned and whose score may be either 1 (i.e. the value is appealed to) or 0 (i.e. the value is not mentioned). The social support is computed according to the affinity between the values highlighted by how the proposal is framed and the values of the policy-subjects. It is assumed that the frame appeals to, at least, one value (i.e. $\exists F_i = 1$).

The *affinity* ξ ranges from 0 (i.e. no affinity at all) to 1 (i.e. perfect affinity). The affinity is calculated as follows (Alg. 3): if the *framing* appeals to a value that the household holds, then they are akin, but if the frame appeals to one value that the household does not hold, then it will look "less ethical". The number of coincidences, normalised by the number of values of the *framing*, determines the *affinity*. Notice that those values that are not appealed to by the *framing* are not checked, as if households do not consider those values that are unmentioned. For instance, a proposal whose *framing* is based on two values has a perfect affinity with a household that holds those two values as well, regardless of the other values that held by the household.

Algorithm 3: Compute the *affinity* between a household and a political demand

```
Input: Framing $\vec{F}$, Household's held-values $\vec{V}$
Output: Affinity $\xi$
foreach pair $\{F_i, V_i\}$ for $i = 1..n$ do
    if $F_i = 1$ then
        $m \leftarrow m + 1$
        if $V_i = 1$ then
            $k \leftarrow k + 1$
return $\xi \leftarrow k/m$
```

For a household to support a demand there is one main condition: the household must be unsatisfied, in which case the household will seek for a *political demand* to support —one that has *affinity* with its values (i.e. is ethically appropriate for the household). In this model, it is considered that households can support only one *political demand*, but it may change each time-step. In this model, all households express themselves politically. Hence, as long as the previous conditions are met, household will support *political demands* (that is, political participation is not optional).

First, if the household's satisfaction is below 0.6 (value set for test purposes), then it will check all active *demands* (i.e. those that have been raised by *policy-influencers* and persist in the social space (Table 10.11); the satisfaction indicator used depends on the value profile of the household:

- *Self-centred* households will use their service-satisfaction (that is, they have to be displeased with the service at home before they seek for political intervention).
- *Conservative* households will use their global-satisfaction (that is, they have to be displeased with both the service at home and the political state before they seek for political intervention).
- *Universalist* households will use their political-satisfaction (that is, they have to be displeased with the political state before they seek for political intervention). This leads to "mass behaviour".

Second, if *affinity* ξ between the demands' *framing* and the households' held-values is greater than or equal to 0.5 (value set for test purposes), then the *political demand* can be supported by the household. Among all the possible *demands*, the household will choose the one with greater *affinity*. In the case that multiple *demands* have the same *affinity*, then the household will choose the one whose *framing* appeals to more values (that is, whose *framing* has more non-null components).

To determine the social support of each *active political demand* (Table 10.11), it is counted the number of households that are supporting it (thus, the social support for one *demand* ranges from 0 to 100 %, and the total social support given to all *demands* is, maximum, 100 %).

Policy-influencers request to enact their political demands

The *municipality* is afforded to enact policy instruments at the institutional level. For this reason, the other *policy-influencers* request the *municipality* to enact their *political demands*, thereby producing policy shifts. This process is implemented as follows: each time-step the *political demands* raised in the social space are checked (Table 10.11). If the *municipality* is afforded to enact them and they have enough social support (see Table 10.14), then they can be enacted. Among these set of possible enactments, the *municipality* will choose the one with greatest social support. After that, the demand is enacted and withdrawn from the social space.

10.5 Simulation

10.5.1 Approaches

The model is simulated by taking two different approaches:

- First, the model is used to simulate simpler scenarios without *policy-influencers* (i.e. "simple use"). Households do not assess the state of affairs at a social level (they focus only on the service) and therefore political demands cannot thrive. This posits a less realistic representation of the world —albeit simpler. The reason is to provide a basic understanding of the outputs to address more complicated scenarios.
- Second, the model is used to simulate scenarios in which there are *policy-influencers* (i.e. "standard use"). Thus, policy shifts can occur, eventually producing different social outcomes. This posits a more realistic representation of the world, but also more laborious modelling and analysis exercises. The simulation of those simple scenarios allows to better understand the social outcome when *policy-influencers* are involved.

The main input for simulation are different value profiles for the population of households. Hence, the model illustrates how different values lead to different social outcomes. For the sake of exaggerating

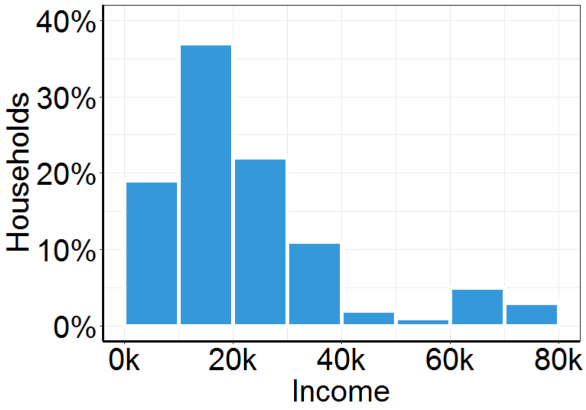


Figure 10.3: Households' income distribution

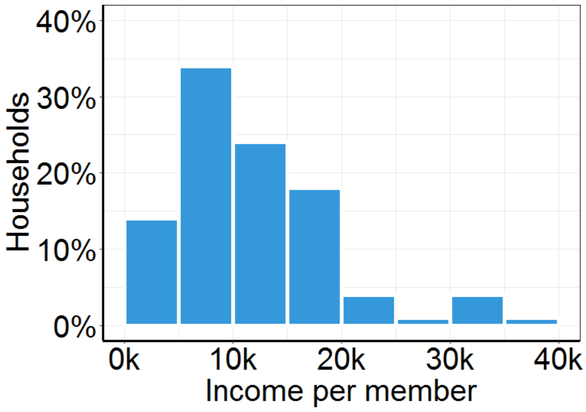


Figure 10.4: Households' income per member distribution

the simulation output, the scenarios represent the outcome when the population is entirely constituted by one type of value-profile. Moreover, for the sake of comparability, households' income is kept constant for the different scenarios (Figs. 10.3 and 10.4).

Policy-influencers start to act after 12 time-steps, in order to give time to households to adopt stable water-use behaviour, in accordance with their value profiles. As long as it is not mentioned explicitly, the default input has both the water supply shutoffs and the social aid disabled.

10.5.2 Simple use

Scenario Sm1

This scenario simulates the social behaviour of a population of households whose *value-profiles* are 100 % *self-centred*. As there is no major events during the simulation, the outcomes are quite stable.

The use of water of this population is intensive (namely around 290 L/p·d on average), since households only adopt conservation habits when the bill is expensive (Fig. 10.5). The service-cost recovery-rate is high (namely, 115 %), due to the high water bills (Fig. 10.6).

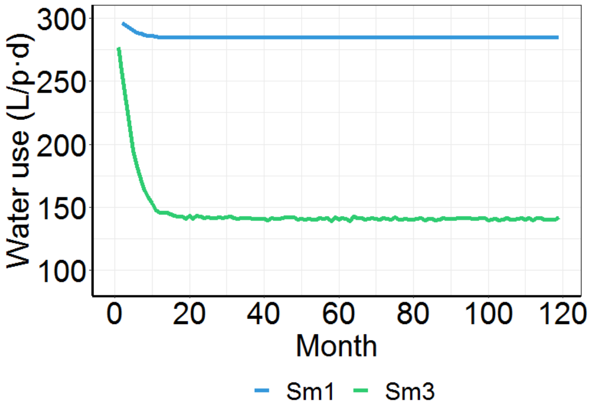
Households are satisfied with the service with regard of the supply (as they can use as much water as they want), but not so much for the bill (the bill is expensive due to their high-use of water). The overall satisfaction of the service is approximately 70 %. Either enabling supply shutoffs or social aid does not change the social outcome.

Scenario Sm2

This scenario simulates the social behaviour of a population of households whose *value-profiles* are 100 % *conservative*. The outcome of this scenario is similar to the Scenario Sm1, in which 100 % of the households are *self-centred*.

Scenario Sm3

This scenario simulates the social behaviour of a population of households whose *value-profiles* are 100 % *universalist*.



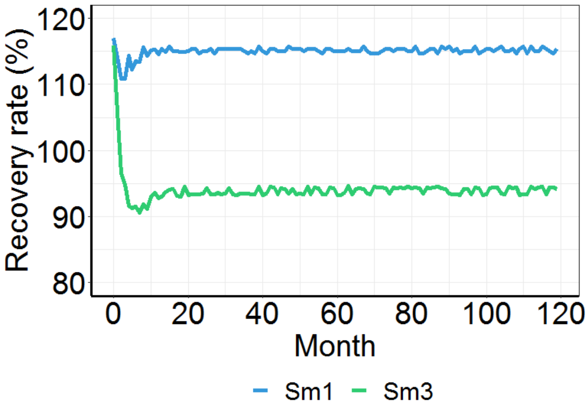


Figure 10.5: Households' average water use depending on the population value-profile

Figure 10.6: Service-cost recovery-rate depending on the population value-profile

In this scenario, water use is much lower, since *universalist* households are motivated to have an environmentally-responsible use besides an affordable bill (the water use is approximately 145 L/p·d on average) (Fig. 10.5). At first, *universalist* households adopt all conservation practices equally. When they have a stable water use around 145 L/p·d, once they get technology (e.g. dishwasher) —which is assumed that it cannot be abandoned once adopted—, this lowers the water bills, inducing them to abandon some conservation habits (e.g. being efficient when flushing the toilet), but without great impact on the average water use. Nevertheless, the overall satisfaction of the service is approximately 90 % (their water use is translated into guaranteed supply and low bills).

Service-cost recovery-rate does not reach 100 %, because households use little water (namely, this indicator is approximately 95 %) (Fig. 10.6).

If the social aid is enabled, the number of households that benefit from it decreases rapidly (due to their reduced water use). In the same vein, the number of households in debt also decreases. In any case, there is a small group of households that, due to their low income, will apply for the social aid on a general basis (as shown in Fig 10.3, there are households whose income is lower than 13,000 eur/year).

10.5.3 Standard use

Scenario St1

This scenario simulates the social behaviour of a population of households, whose *value-profiles* are 100 % *self-centred*, and the multiple *policy-influencers*.

Due to the high use of water (approximately 290 L/p·d), the *municipality*, whose main value is *water security*, enables a maximum supply for households of 120 L/p·d, similar to the instruments that have been considered in cities of Australia or Cape Town (see [241, 266]) —the difference is that, in this model, it is not possible to not abide by this norm) (Fig. 10.7). This causes a drop in the water use (i.e. 120 L/p·d), that is quite different to the actual water demand (i.e. 290 L/p·d), which translates into a very low valuation of the service, even though they are satisfied regarding the water bill. It also entails lower water bills, which reduces the service-cost recovery-rate from 115 % to 95 % (Fig. 10.8). Under these new circumstances, households adapt their water use. They start to adopt conservation practices during the year after the enactment of the maximum supply. Besides, since *water security* is guaranteed, and the service-cost recovery-rate is acceptable, the *municipality* compensates households with lower rates —which causes a reduction in the service-cost recovery-rate to 85 % (Fig. 10.8) and, consequently, also in the households' political satisfaction (*competence* value of trusted *political-influencers*) (Fig. 10.9).

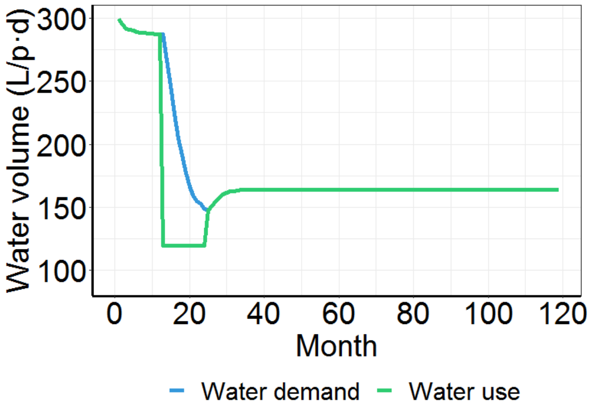


Figure 10.7: Households' average water use and average water demand in the scenario St1

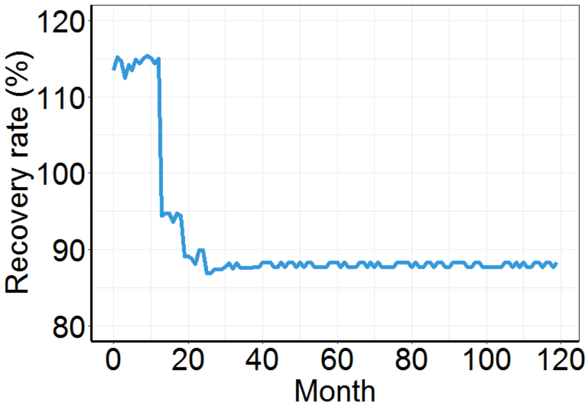


Figure 10.8: Service-cost recovery-rate in the scenario St1

During this stage the *municipality* is monitoring the *water demand* of households. When they have adapted their water-use behaviour so to demand approximately 150 L/p·d, the *municipality* disables the maximum supply (Fig. 10.7). Since the households have low water bills, this causes them to abandon some conservation practices, but they keep their average water use within a level acceptable by the *municipality* (from 120 L/p·d to 165 L/p·d). This causes an increase in the service satisfaction of households, because their water demand is completely satisfied (Fig. 10.9).

Meanwhile, the *water utility* advocates for raising water fees to ensure the financial sustainability of the service —nevertheless, its political satisfaction is pretty high because households' water needs are met with an acceptable amount of water possible.

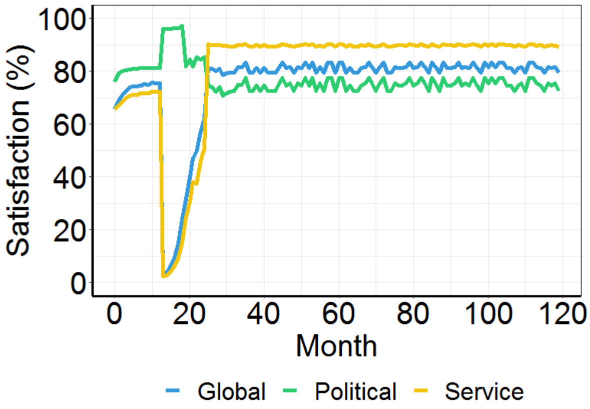


Figure 10.9: Households' average satisfaction in the scenario St1

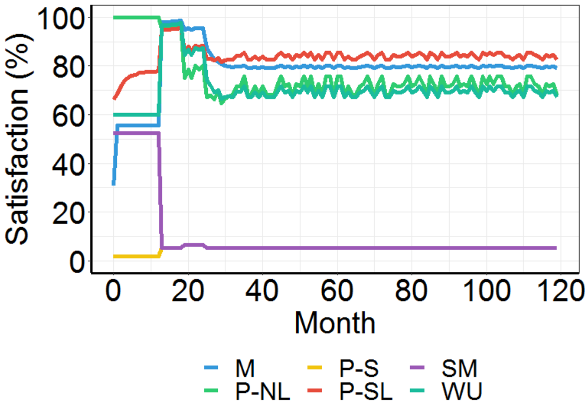


Figure 10.10: Policy-influencers' satisfaction in the scenario St1

The *SL-party*, despite being less satisfied with respect to *competence*, it is quite pleased with regard to *social justice*: on average, water bills are reduced, which means that the service is affordable from its point of view. The *NL-party* starts from with complete satisfaction (because the service-cost recovery-rate is high), but it drops after the *municipality* intervenes (Fig. 10.10). In some simulations, this may lead it to demand to increase the private participation in the water management, since the service *is not competent*. If this occurs during the use-restriction, all households support that demand (as they are completely displeased with the service), forcing the *municipality* to execute that measure. After this happens, the *NL-party* is more "benevolent" when evaluating the *competence*, increasing its political satisfaction despite the service-cost recovery-rate not changing.

The *social movement* and *S-party* are completely dissatisfied (Fig. 10.10): *social justice* and *control* issues are not addressed directly. Motivated by this, they advocate for direct management and social mobilisation, respectively, without having any success.

If water shutoffs are allowed, the *S-party* and the *social movement* increase their satisfaction with regard to *social justice*: in the way their *value profile* is modelled, when shutoffs are allowed, these *policy-influencers* no longer focus on the impact water fees have on vulnerable households, but instead they focus on the number of households that have their supply interrupted. As the *municipality* enacts a maximum supply instrument, practically all households are able to pay the bills, so there is only a few households that face a water shutoff. If the social aid is enabled, the *NL-party* gathers enough social support to disable such instruments, before demanding greater private participation in the water management (therefore, this instrument does not produce any effect).

Scenario St2

This scenario simulates the social behaviour of a population of households whose *value-profiles* are 100 % *conservative* and the *policy-influencers*.

The outcome of this scenario is similar to the previous case. The main differences lay on the "political successess" of the *policy-influencers*. For instance, if the *NL-party* demands for more private participation in the water management, in most of the simulations this demand is not successful.

During the water restriction period, the *water utility* is able to maintain the *water fees* (that is, it counters the attempts of the *municipality* to lower the water fees), since this measure is framed as an intervention to promote *competence*, *efficiency* (i.e. a more "rational" use of water) and *water security* (e.g. "using little water today means having more water for tomorrow"). After the restriction is lift, the rebound effect in households' water use is not produced (in comparison to what happens in Fig. 10.7).

Scenario St3

This scenario simulates the social behaviour of a population of 100 % *universalist* households and the *policy-influencers*.

In the first year, households achieve a water use of 140 L/p·d (without intervention of the *municipality*), which causes the service-cost recovery-rate to drop to 90 % (see Scenario Sm1). Once *policy-influencers* start to interact with households, one of the first measures to be enacted is to increase the public participation of the water management, raised by the *S-party*. Since all households are *universalist*, and this measure is framed with the values of *control*, *social justice* and *citizen participation*, the proposal is immediately executed. Meanwhile, the *NL-party* advocates for increasing the private participation, unsuccessfully.

The *social movement* focuses on *social justice*, namely on the impact of water fees on vulnerable households. According to the input statistical data, there is a group of households with low income, that makes them particularly vulnerable and sensitive to water fees (Fig. 10.3). Thus, since the evaluation is focused on this disadvantaged group, the *social movement* immediately advocates for social mobilisation, in order to not pay the water bills as a measure of social protest. Consequently, all the households support the measure —after having reduced private participation in the water management— and declare themselves insubordinate. This causes the service-cost recovery-rate to drop to 0 % (since the protest is 100 % supported) (Fig. 10.11).

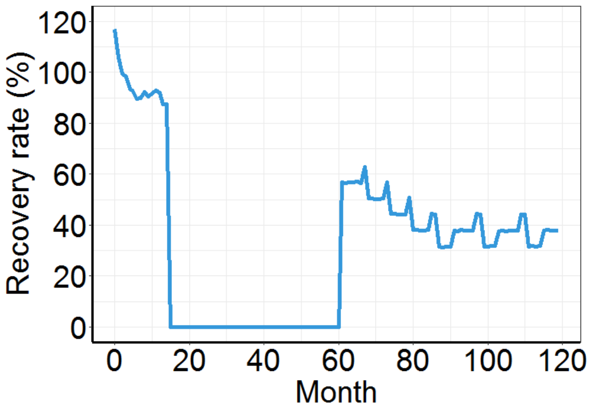


Figure 10.11: Service-cost recovery-rate in the scenario St3

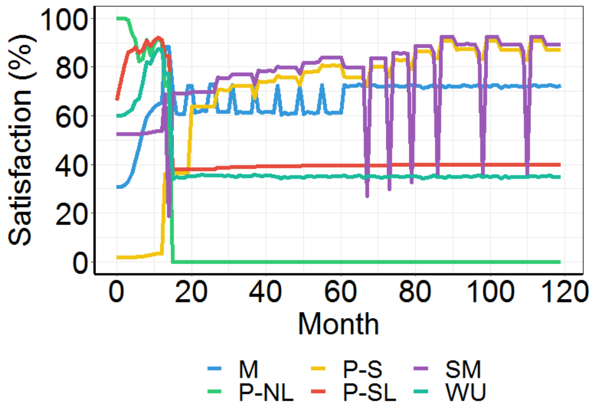


Figure 10.12: Policy-influencers' satisfaction in the scenario St3

The *water utility* and the *SL-party* advocate for increasing water fees to ensure the service-cost recovery, but they cannot gather any social support. The *municipality*, that is afforded to increase of water fees regardless of their social support, does so in order to deal with the drop of recovery of service costs. Despite the *municipality* trying to raise water fees, both the *social movement* and the *S-party*, backed by the entire population of households, manage to decrease the water fees to very low levels (with a service-cost recovery-rate of approximately 35 %). After that, they call the social mobilisation to finish. Besides, at some point after the social mobilisation has been started, the *social movement* achieves the enactment of the social aid. Anyway, in the end, practically no household applies for the social aid due to the low fees.

According to the way value profiles are implemented, the *social movement* experiences high variation in its political satisfaction (Fig. 10.12). The reason is that, after it achieves to push for its demands, it is highly satisfied with regard to *citizen involvement* (i.e. all households have supported one of its demands), which causes its political satisfaction to rise. This causes that households are also politically satisfied, to the point they will not support for the coming *social movement* political demands (that is, they consider that the world-state is fine, because affine *policy-influencers* —and especially the *social movement*— said so, as they were successful in pushing their demands). Thus, when *social movement* advocates for *social justice* related demands (e.g. reduce water fees, social aid, etc.), since households are satisfied, the *social movement* does not gather any social support for those demands. This causes the *social movement* to be displeased with regard to *citizen involvement*, which is communicated to households to decrease their political satisfaction and make them open to consider political interventions ("the world-state is not fine yet and nobody is doing anything to address it!"). In some cases, it is cyclic: (1) households do not support the *political demand*; (2) the *social movement's* satisfaction with regard to *citizen involvement* decreases; (3) then households are displeased and support the *political demand*; (4) making the *social movement's* satisfied with regard to *citizen involvement*; and then back to (1).

According to this, when the number of households in debt is reduced, the satisfaction with regard to *social justice* of *policy-influencers* increases, which may make households so satisfied to the extent they will not seek any political demand to support. This causes satisfaction with regard to *citizen involvement* to drop because households do not support *social movement's* demands. The same can happen when increasing the share of public participation in the water utility. Thus, ironically, improvements in *social justice* and *control* may lead to social discontent, according to the model.

If social aid is enabled, the outcome is practically the same. The *NL-party* advocates for suppressing this aid without any success. If water shutoffs are enabled, the outcome does not change notably:

prior to the call to social mobilisation, the *social movement* advocates for banning such practice.

When these two measures are combined (social aid and water shutoffs), the outcome is, in general, a stable case without social mobilisation (Figs. 10.13 and 10.14). It happens when vulnerable households are able to deal with their debts in the first time-steps (remember that it is partially random), while the *social movement* and the *S-party* are demanding an entirely-public operated management (which successes). After this, as vulnerable households have managed to escape from this vicious circle of debts with the help of the social aid, there are no water shutoffs, despite being allowed —which is what the *social movement* and the *S-party* are monitoring to evaluate the world-state with regard to *social justice*. Consequently, the *social movement* is satisfied (as there are no shutoffs) and does not have any political demand. Households are also satisfied, since they are informed through the social movement, so they are not motivated to support any demand. Meanwhile, the *S-party* advocates for disabling the water shutoffs, but it does not gather social support because households are already satisfied.

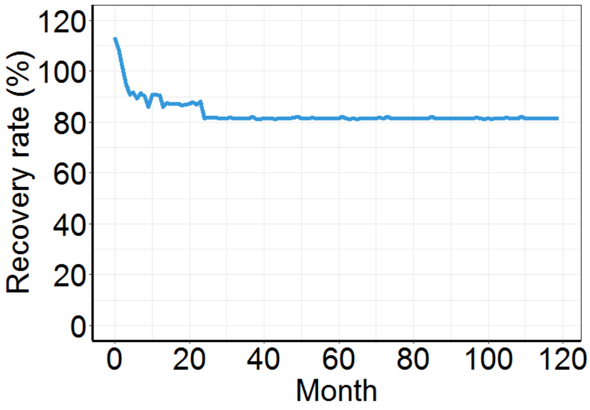


Figure 10.13: Service-cost recovery-rate in the scenario St3 with *water shutoffs* and *social aid* enabled

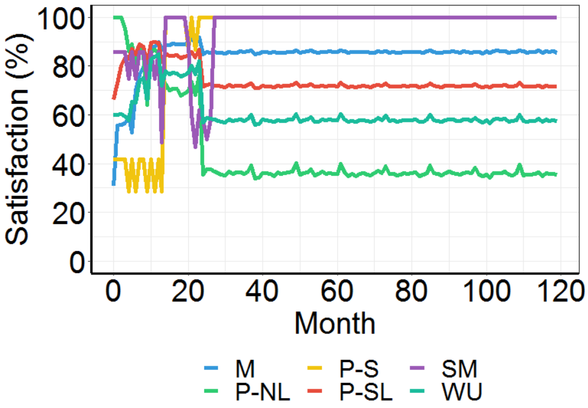


Figure 10.14: Policy-influencers' satisfaction in the scenario St3 with *water shutoffs* and *social aid* enabled

Scenario St4

This scenario simulates the social behaviour of the *policy-influencers* and a population of households whose *value-profiles* are 50 % *universalist*, 25 % *self-centred*, and 25 % *conservative* (as shown in scenarios Sm1 and Sm2, *self-centred* and *conservative* lead to similar outcomes). Additionally, (1) households' income has been reduced to 25 % (economic crisis), and (2) *social aid* and *water shutoffs* are enabled from the beginning.

The average water use drops during the first year, from 320 L/p·d at the start of the simulation to 140 L/p·d (Fig. 10.15). Notably, many households have debts with the water utility —number that decreases with the level of water use. Most of them are beneficiaries of the social aid. On average, 10 % of the households suffer from water shutoffs (Fig. 10.17), which causes that water demand is greater than water use on average (Fig. 10.15). The service-cost recovery-rate is sensitive to this: when households struggle to pay the water bills, the indicator fluctuates rapidly (approximately from 90 % to 75 %) (Fig. 10.16), which impacts on the political satisfaction of the *policy-influencers* that hold the *competence* value (and, consequently, on *conservative* and *self-centred* households' satisfaction (Fig. 10.18)).

The *municipality* practically does not intervene, and, if it does, it reduces the water fees. The *water utility* tries to increase the water fees, to achieve appropriate service-cost recovery. Meanwhile, the

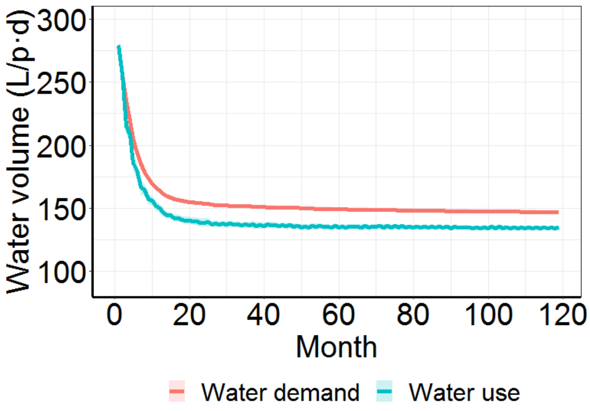


Figure 10.15: Service-cost recovery-rate in the scenario St4

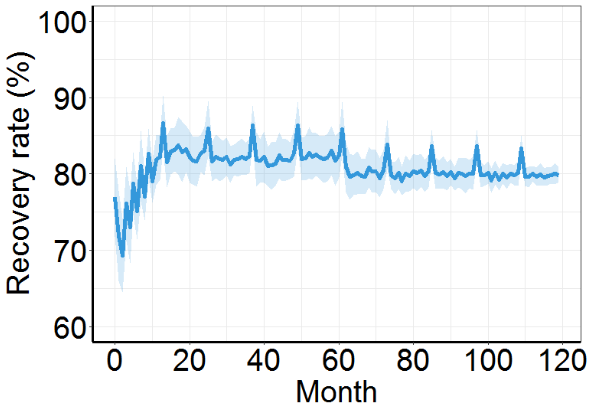


Figure 10.16: Policy-influencers' satisfaction in the scenario St4

S-party advocates for disabling water shutoffs. However, *universalist* households support social mobilisation raised by the *social movement*.

Nonetheless, social mobilisation does not succeed, as there are not enough households that support it —the social support reaches exactly the 50 %, while the action requires **more than** 50 %. Other households support the increase of tariffs (raised by the *water utility*), while others want more private investment to improve *competence* (raised by the *NL-party*). Anyway, households are in general unsatisfied (Fig. 10.18), which indicates that the situation is socially unstable.

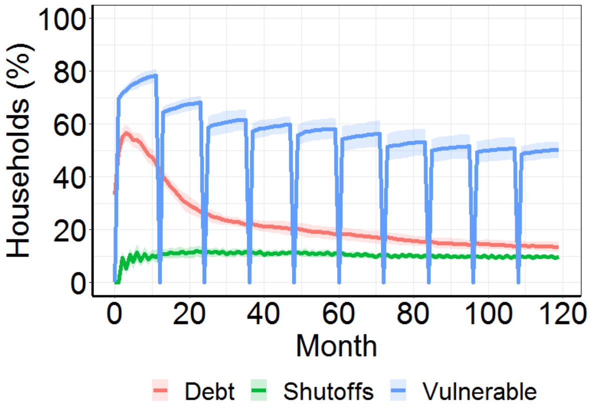


Figure 10.17: Vulnerable households in the scenario St4. Those households that benefit from social aid are more than half of the population —notice that each year the social aid is reset, and households must apply for it again.

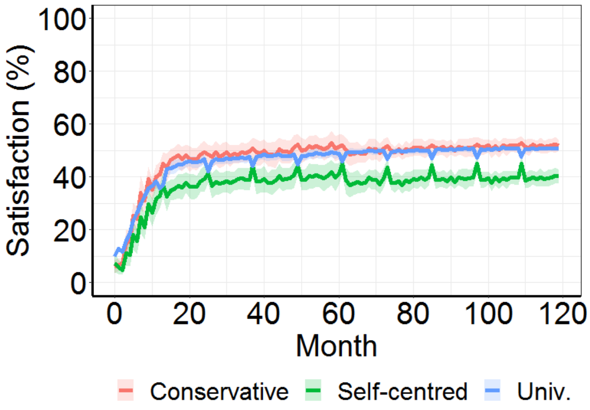


Figure 10.18: Households' global satisfaction depending on their value-profile in the scenario St4

10.6 Discussion

10.6.1 Insights withdrawn from the simulations

Although the model is not validated with real data, the simulations show realistic outcomes. First, unsatisfied social expectations concerning values may lead to social acceptance issues and political shifts, as indicated in other works (see [269], which covers the energy domain). Second, the exploration of outcomes may *teach* interesting conclusions about the domain.

The simulation suggests that agents concerning only wealth-related values (e.g. *competence*, *wealth*) may be a barrier to *efficient* water uses (i.e. 120 L/p·d) (Scenario Sm1). The water prices are, in general, *affordable* for half of the population when they use high levels of water (i.e. 275 L/p·d), since the median of the relative water bill (i.e. the ratio between the water bill and the household's income) is approximately 3.6 % (according to the UN, the service is *affordable* if the cost is lower than 3 % the household's income [257]). As these agents only react to economic signals, increasing water fees to reduce water use pushes most of the population to economic difficulties. Thus, the problem is to progressively decrease water use without causing economic difficulties, but using only economic instruments (because households do not respond to other values). Another option is to rationing supply, which is *effective* to achieve an *efficient* water use, but it also leads to low levels of social acceptance (Figs. 10.7 and 10.9). According to the present model, low levels of social acceptance *trigger* political shifts. Hence, *policy-influencers* could take advantage of the social unrest to promote extreme interventions (this is why it is interesting to consider dynamic value profiles in future works).

Adopting conservation practices contributes to reduce their water bill of vulnerable households. Yet, this should not be the strategy of the authorities and water managers to avoid responsibility concerning *affordability* for these households. Many conservation practices are afforded by household appliances (e.g. eco-mode programmes of washing machines), and it must be expected that vulnerable households cannot actually adopt them (thus making them unable to reduce significantly water bills by means of adopting conservation practices). Furthermore, as shown in Scenario St4, providing vulnerable households financial support may contribute positively to the service-cost recovery rate (in plain terms: a household that pays little contributes more than a household that pays nothing) (Figs. 10.16 and 10.17). However, this instrument is specially subject to the notions of *fairness* of the *policy-influencers* and households (i.e. they may consider that a household paying less than another household, no matter its situation, is unfair and therefore it should not be afforded). Thus, it is advisory for these kind of counter-intuitive instruments to check the values of the population and *policy-influencers* to decide whether it is necessary to launch a communication campaign to debunk potential myths.

One clearly revealed value conflict is the one between *competence* (i.e. financial sustainability of the service) and *social justice* (i.e. affordability of the water service for every household). Trying to achieve *competence* by means of increasing water fees will impact negatively on *social justice*. Likewise, *competence* conflicts with *efficiency* (i.e. prudent use of natural resources), because the less water is used, the fewer the resources are collected by using water fees.

For this same reason, satisfaction of *policy-influencers*, concerning different values, might be directly opposed (e.g. *social movement* and *NL-party*). This means that policies for water public services are committed to certain values while leaving others unattended, thus causing "winners" and "losers". The challenge here is to define the problem of management to consider the multiple values involved and justify decisions according to this view (that is, not neglecting some values, but incorporating them within the managed reality in order to generate other kind of solutions).

10.6.2 Simulation of political phenomena

As explained for the case study in Chapter 9, it is a strong assumption to represent satisfaction as a numerical measurement and use this as a *causal mechanisms* of the simulation model [71]. Social acceptance and political action may be motivated (or rejected) by diverse values and multiple causes (see [269]). For this reason, political behaviour arise from a more complex model of rationality, which points for further research in how to model it in computational models.

Consequently, simulators to explore such phenomena are extremely difficult to model and validate. Since the socio-political world is complex, achieving to appropriately abstract policy-influencers behaviour may be impracticable: participating in political arenas involves many resources (e.g. economic, cognitive, etc.), multiple political strategies and social interactions (e.g. long-term negotiation, symbols disputation), and diverse political variables (e.g. power, legitimacy, etc.). Nonetheless, although these models cannot be used for prediction policy outcomes, they may be useful for other purposes (see [73]).

Thus, simulators grounded on the knowledge of domain experts may be helpful to test their intuitions. The very exercise to model the value profiles of stakeholders may clarify what they value, which may lead to adopt additional instruments to prevent, mitigate, or compensate damages. Hence, the simulator contributes to explore in a more formal and analytical approach the effects of interventions: as long as the model is realistic, it may inform the design of interventions. For public service management, it can be useful for monitoring users' satisfaction under different scenarios (either policies or events), and it can be more powerful if combined with hydraulic models.

10.6.3 Policy-assessment support system

Despite not being implemented in the model presented in this chapter, a policy-decision support system would have an additional interface layer between the user (i.e. policy-maker) and the simulation. Because the intention of this chapter is to illustrate the stakeholders' value-driven behaviour and computational approaches to simulate relevant value-related aspects, it has not been implemented so as not to commit nor imply a particular policy decision-support framework (i.e. relevant values and indicators, hierarchies, assessment, etc.). Anyway, it is convenient to mention that value-driven simulators for policy assessment would rely on such frameworks to support the choice between alternative options.

The decision-support layer would define the concrete framework by which the policy-maker is assessing the (simulated) policy outcomes, eventually guiding its decision (See Sec. 4.6 in Chapter 4). That is, the policy-maker would establish a set of values and indicators corresponding to the values intended to be promoted or preserved, a set of evaluation/aggregation mechanisms —as been illustrated for the *simulated* stakeholders (see Appendix A)— and an assessment scope (e.g. time horizon). Among the most basic decision-support outputs there could be a value assessment matrix (alternative policies against their assessment corresponding to different values) and a overall score for each policy option .

Interestingly, the model *limits* the decision-support framework, and this is an ethical concern for model designers. In other words, the policy-maker may not be able to "assess" the policy outcomes with regard of some values unless these are not part of the model. Therefore, the very model may bias the decision-support frameworks.

10.6.4 Further work

In the simulation only few policy shifts can be observed. For instance, in the scenario St4, the *social movement* intends to carry out a drastic shift (in the sense that it entails significant effects but requires much social support), instead of minor interventions, thus keeping the situation “unchanged”. This points out to the necessity to expand action sets of *policy-influencers* and improve their capability of reading the situation appropriately (that is, adapting their *political demands* to minor interventions if they do not gather enough social support).

It is expected that a more sophisticated social model could make simulations more realistic (new actions for *policy-influencers*, such as ‘political-attacks’, ‘negotiation’, etc., and new actions for *policy-targets* like ‘imitation’, etc.). Likewise, further research should address how to build realistic social networks in urban environments.

Furthermore, in the present model, *policy-influencers* have the same capacity and legitimacy to push for their demands, but that may be not the case in reality (for instance, some administrations may be more open to negotiate with right-wing political parties rather than leftist social movements). This lead to consider the financial behaviour of the water-management company (involving profits, taxes, reinvestment, subsidiary companies, corruption scandals, etc.), since it is extremely relevant for the political world. Presumably, it is a source of social conflict, as agents may perceive the company’s behaviour as unfair while their behaviour is being increasingly regulated. For instance, it is a common argument for social movements to point out that the water utility is reporting profits while households struggle to pay the water bill.

Besides, notice that this model constitutes a particular understanding of the management of urban water. In other words, the main *policy-target* are households, but other entities could be also relevant: municipal water users, urban industries, etc. Aside from this, other simulators could use the notion of *policy-influencer* not as actual political actors, but as stakeholders in a management board within the utility company.

Finally, at this level of work, the modeller can only speculate about how values define households’ behaviour in the urban water domain (thus, simulations are purely hypothetical experiments). Noteworthy, Schwartz *et al.* explored the relation of core values to voting in Italian elections [227]. Therefore, election results could guide the inference of value profiles, but it should be appropriately studied in the Spanish case. That said, more fieldwork is necessary to explore values and behaviours (not only profiles) to build more realistic agents. Also, although values profiles are said to be stable during adulthood [183], it may be interesting to consider that their understanding and evaluation mechanisms can change due to political messages in the social space.

10.7 Closing remarks

A prototype based on agent-based simulation to explore the role of values in social acceptance and policy shifts in the urban water realm has been presented in this chapter. Satisfaction and political action have been modelled using the notion of values: they express social expectations that are assessed by observing and evaluating particular components of the *reality*. More precisely, *policy-influencers* communicate evaluations concerning the whole systems in order to influence households to make them to subscribe their political views and support their proposals. Accordingly, different value profiles for the population of households lead to different social outcomes.

The model has been built as an hypothetical situation to explore the social outcomes and the social acceptance under multiple scenarios concerning different population value profiles. The applicability of the simulation of these models is quite promising, since it could be used to improve management policies and social acceptability in different challenging situations (e.g. heat waves, economic slowdown, etc.).

However, although the model is realistic, simulation of these phenomena requires much field work (i.e. values, behaviour models, actions, social interactions, etc.). Many disciplines are involved and, consequently, much effort has to be done to effectively assemble diverse knowledge from these fields (e.g. psychology, economy, water engineering, sociology, etc.). Working on particular case studies would allow a more accurate representation of the actual phenomena that is explored, besides including concrete insights from experts. At this point, this work can only present a simple prototype that illustrates how these simulators may be useful as policy-support systems.

10.8 Appendix A: Policy-influencers

10.8.1 Municipality

Values. Its core values are *control*, *competence*, *citizen involvement*, and *water security* (Table 10.12). This stakeholder considers that its main task is to monitor and address *water security* to protect the population from water-related threats. Moreover, it has to ensure that citizens can have a say in how these public services are managed (i.e. *citizen involvement*). It is convenient that these water services are financially sustainable and provide an appropriate service to citizens, as much as possible (i.e. *competence*). Finally, since the municipality is the holder of- and ultimate responsible for the water public services, this stakeholder considers that *control* to be relevant —which could facilitate their tasks.

Table 10.12: Values of the municipality

Value items (ranked)	**Weight**	**Main indicators**
Water security	4	Average water use (L/p·d)
Competence	2	Service-cost recovery rate (%)
Control	2	Public participation in the water utility (%)
Citizen involvement	1	Households supporting a *political demand* (%)

Value functions:

- **Water security.** This function assesses the world-state with respect to *water security* using the average water use (namely, the average revenue water, in L/p·d) as indicator (Fig. 10.19).
- **Competence.** This function assesses *competence* of the management using the service-cost recovery rate through water fees (in %) (Fig. 10.20).
- **Control**. This function evaluates the world-state using the share of public participation in the water utility (in %) (Fig. 10.21). It is only executed if the average *political satisfaction* of households is lower than 0.6 (that is, *control* is not an important issue unless households are politically displeased, which will require political intervention).

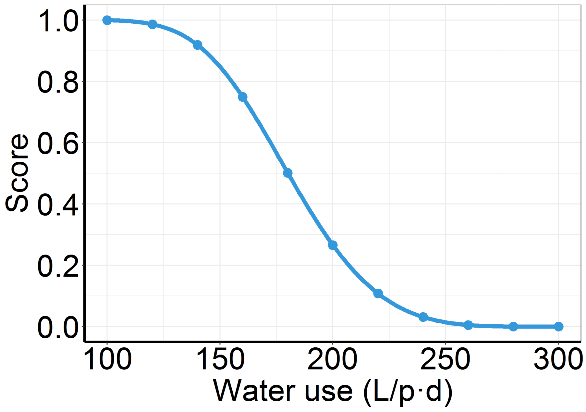


Figure 10.19: Municipality's value function for *water security*

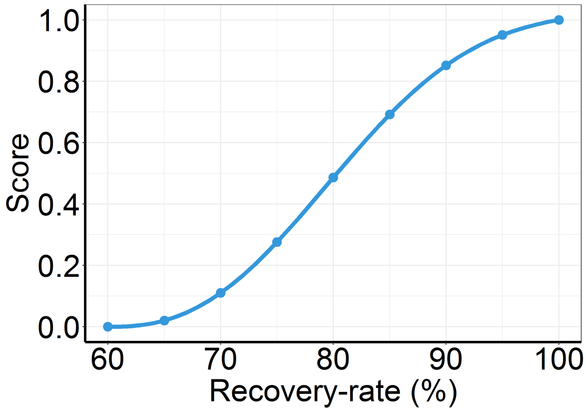


Figure 10.20: Municipality's value function for *competence*

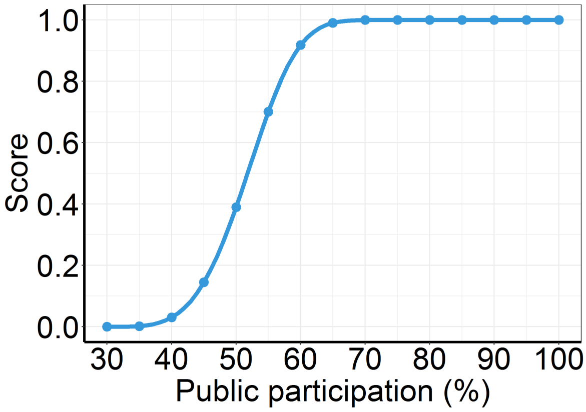


Figure 10.21: Municipality's value function for *control*

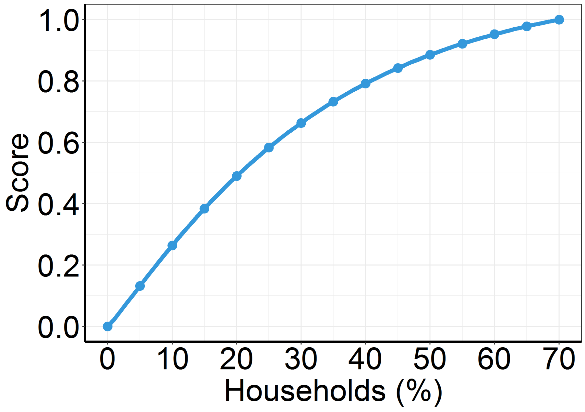


Figure 10.22: Municipality's value function for *citizens involvement*

- **Citizens involvement**. This function assesses the *citizens involvement* using the number of households that support a *political demand* (in %) (Fig. 10.22). It is only executed if the average *political satisfaction* of households is lower than 0.5 (that is, if households are politically displeased, the *municipality* considers that they should express their political views).

Political demands. Its main actions are intended to maintain the *water security*, by means of economic instruments (e.g. water tariffs), normative instruments (e.g. establishing a maximum supply) and persuading instruments (i.e. campaign to increase environmental awareness) (Table 10.13). Moreover, some of these instruments may be used to increase *competence*, in terms of covering the service costs. As long as these values are ensured, these means may be relaxed.

Political requests. The *municipality* can enact some *political demands* of other *policy-influencers*, as long as these gather enough social support from households (Table 10.14).

Table 10.13: Set of political demands of the municipality

Demands	Preconditions	Additional (adv.)	Additional (enact.)	Framing	Type
Enable maximum supply	$C_6 < 0.1$	Maximum supply is disabled	—	—	Enactable
Disable maximum supply	$C_6 > 0.8$	Maximum supply is enabled The average *water-demand* of a 10 % of the households, randomly selected, is below 150 L/p·d	—	—	Enactable
Increase tariffs	($C_2 < 0.6$ and $C_6 < 0.4$) or ($C_2 < 0.3$)	—	—	—	Enactable
Reduce tariffs	($C_2 > 0.8$) and ($C_6 > 0.8$)	$\gamma > 0$	—	—	Enactable
Environmental awareness campaign	$C_6 < 0.6$	—	—	—	Enactable

Table 10.14: Set of requests sent to the municipality

Requests	Preconditions	Additional
Enable supply shut off	Social-support > 60 %	Supply shut off is disabled
Disable supply shut off	Social-support > 20 %	Supply shut off is enabled
Increase tariffs	Social-support > 40 %	—
Reduce tariffs	Social-support > 40 %	$\gamma > 0$
Enable social aid	Social-support > 30 %	Social aid is disabled
Disable social aid	Social-support > 30 %	Social aid is enabled
Direct management	Social-support > 40 %	Public-participation < 100 %
Increase private participation	Social-support > 40 %	Public-participation > 0 %

10.8.2 Water utility

Values. Its core values are *competence* and *efficiency* (Table 10.15). This stakeholder considers its main task to supply water economically and offer a high-quality service (i.e. *competence*). Additionally, since the public service uses water resources in order to meet citizens' needs (namely, supply water for multiple essential domestic purposes like hygiene and sanitation), it has to ensure that this objective is met with the least amount of water resources (i.e. *efficiency*).

Table 10.15: Values of the water utility

Value items (ranked)	Weight	Main indicators
Competence	3	Service-cost recovery rate (%)
Efficiency	2	Average water use (L/p·d)

Value functions:

- **Competence.** This function assesses *competence* of the management using the service-cost recovery rate through water fees (in %) (Fig. 10.23).
- **Efficiency.** This function assesses the *efficiency* of the public service as the average water use (in L/p·d) to met the citizens' needs (Fig. 10.24).

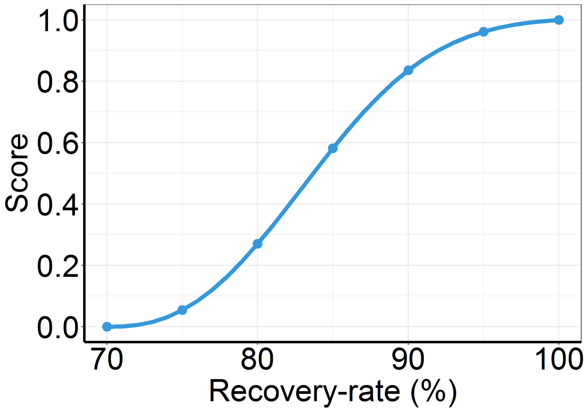


Figure 10.23: Water utility's value function for *competence*

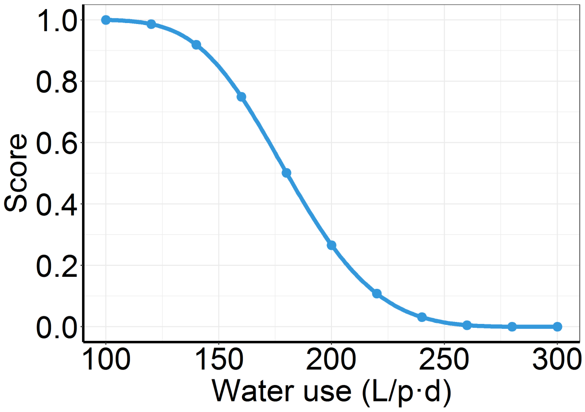


Figure 10.24: Water utility's value function for *efficiency*

Political demands. Its main actions are intended to increase the *competence* by means of economic instruments (e.g. water tariff) and increase the *efficiency* by means of persuading instruments (e.g. campaign to increase environmental awareness) (Table 10.16).

Table 10.16: Set of political demands of the water utility

Demands	**Preconditions**	**Additional (adv.)**	**Additional (enact.)**	**Framing**	**Type**
Increase tariffs	$C_2 < 0.8$	—	—	$F_2 = F_3 = F_6 = 1$	Advocated
Environmental awareness campaign	$C_3 < 0.8$	—	—	—	Enactable

10.8.3 Social movement

Values. Its core values are *control*, *social justice*, and *citizen involvement* (Table 10.17). This stakeholder advocates for the municipality to have the effective *control* of the public service. Moreover, citizens should participate in expressing their opinions with regard to the water governance (i.e. *citizen involvement*). The main focus is to have means to foster *social justice*, namely protecting the rights of citizens and ensuring a safe and affordable water supply for everybody.

Table 10.17: Values of the social movement

Value items (ranked)	**Weight**	**Main indicators**
Citizen involvement	3	Households supporting an affine *political demand* (%)
Social justice	2	Households with their supply shut off (%); Average share of the income that is used to cover the water bill of those households whose income is lower than 13,000 eur (%)
Control	1	Public participation in the water utility (%)

Value functions:

- **Citizens involvement**. This function assesses the *citizens involvement* using the number of households that support a *political demand* that the *social movement* would advocate for (see Table 10.18) (in %) (Fig. 10.25). When the *social movement* is not advocating for any intervention, or it has already managed to enact it, its satisfaction is 1.
- **Social justice.** The *social movement* assesses the world-state with respect to *social justice* using two different value functions. If supply shutoffs are enabled, *social justice* is assessed by means of the number of households whose supply has been shut off (in %) (Fig. 10.27). If supply shutoffs are disabled, *social justice* is assessed by means of the average water bill relative to the income, for those households whose income is lower than 13,000 euros per year (in %) (Fig. 10.28).
- **Control**. This function evaluates the world-state using the share of public participation in the water utility (in %) (Fig. 10.26).

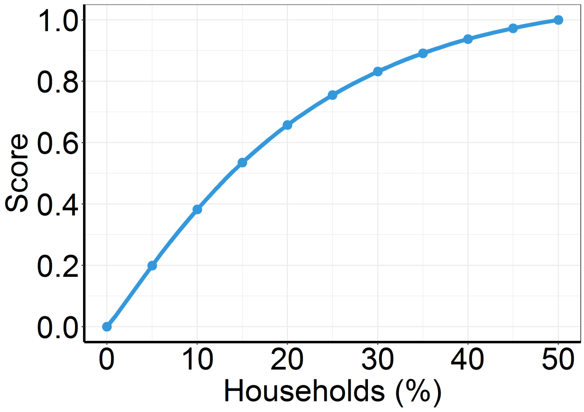


Figure 10.25: Social movement's value function for *citizen involvement*

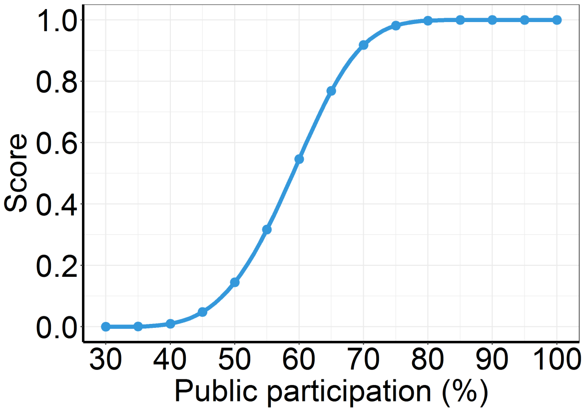


Figure 10.26: Social movement's value function for *control*

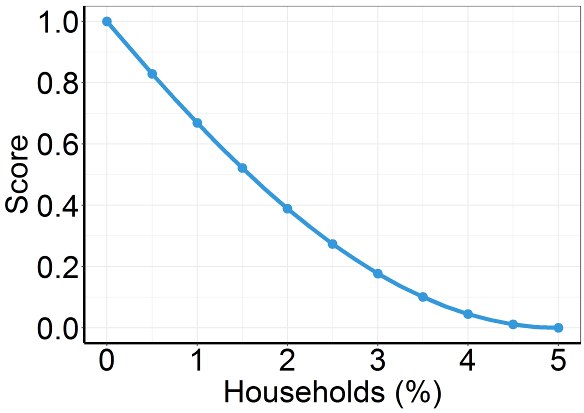


Figure 10.27: Social movement's value function for *social justice* if water-supply shutoffs are enabled

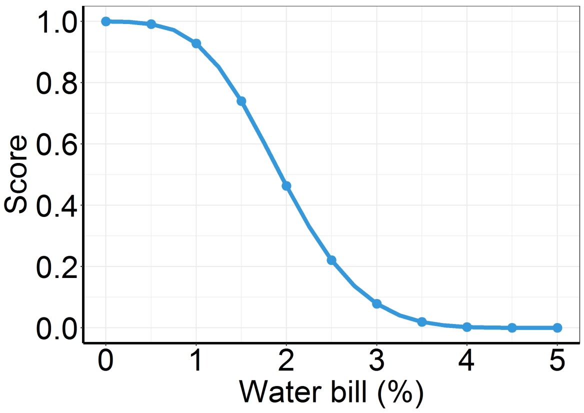


Figure 10.28: Social movement's value function for *social justice* if water-supply shutoffs are disabled

Political demands. Its demands are intended to ensure the *social justice*, by means of economic instruments (e.g. reducing water tariffs), normative instruments (e.g. water shutoff ban), financial instruments (e.g. social aid), and persuading instruments (e.g. social coordination to not pay water bills massively) (Table 10.18).

Table 10.18: Set of political demands of the social movement

Demands	**Preconditions**	**Additional (adv.)**	**Additional (enact.)**	**Framing**	**Type**
Social mobilisation	$C_4 < 0.2$	Mobilised households < 50 %	Social-support > 50 %	$F_4 = F_5 = 1$	Advocated + Enactable
End mobilisation	$C_4 > 0.5$	Mobilised households > 0 %.	—	—	Enactable
Direct management	$(C_1 < 0.5)$ or $(C_4 < 0.6)$	Public-participation < 100 %	—	$F_1 = F_2 = F_4 = 1$	Advocated
Disable supply shut off	$C_2 < 0.8$	Supply shut off is enabled.	—	$F_4 = 1$	Advocated
Enable social aid	$C_4 < 0.8$	Social aid is disabled.	—	$F_4 = 1$	Advocated
Reduce tariffs	$C_4 < 0.6$	$\gamma > 0$	—	$F_4 = 1$	Advocated

10.8.4 Political party “NL”

Values. Its core value is *competence* (Table 10.19), understanding that the public service must not entail excessive costs for citizens. It claims that, for this to happen, *competent* management is necessary, which private companies are more likely to achieve in comparison to public entities, since they have more resources and more incentives.

Table 10.19: Values of the political party “NL”

Value items (ranked)	**Weight**	**Main indicators**
Competence	1	Service-cost recovery-rate (%)

Value functions:

- **Competence**. This function evaluates the *competence* of the management using the service-cost recovery-rate with water fees (in %) (Figs. 10.29). If the share of public participation in the water utility is larger than the private’s, the value function is slightly more strict.

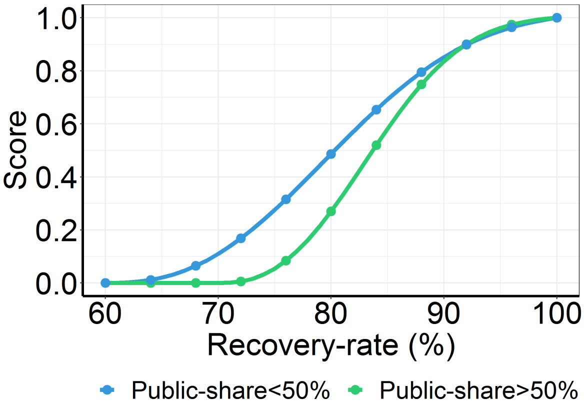


Figure 10.29: NL-Political party’s value functions for *competence*

Political demands. Its demands are intended to ensure the *competence*, by means of suppressing financial instruments oriented to social assistance (e.g. disable social aid), and increasing the participation of private stakeholders in the water utility (Table 10.20).

Table 10.20: Set of the political demands of political party “NL”

Demands	**Preconditions**	**Additional (adv.)**	**Additional (enact.)**	**Framing**	**Type**
Private participation	$C_2 < 0.5$	Public-participation > 0 %	—	$F_2 = 1$	Advocated
Disable social aid	$C_2 < 0.8$	Social aid is enabled.	—	$F_2 = 1$	Advocated

10.8.5 Political party "SL"

Values. Its core values are *competence* and *social justice* (Table 10.21). This stakeholder advocates for a public service that protects citizens' interests (e.g. affordable, universal access, etc.) but it has to be economically sustainable (i.e. *competence*) —otherwise, it would not be resilient enough (to political attacks or economic shocks, for instance).

Table 10.21: Values of the political party "SL"

Value items (ranked)	Weight	Main indicators
Competence	3	Service-cost recovery rate (%)
Social justice	2	Average share of the income that is used to cover the water bill of households (%)

Value functions:

- **Competence**. This function evaluates the *competence* of the management using the service-cost recovery-rate by means of water fees (in %) (Figs. 10.30).
- **Social justice.** This function assesses the world-state with respect to *social justice* using the average water bill relative to the income, taking all the households into account (in %) (Fig. 10.31).

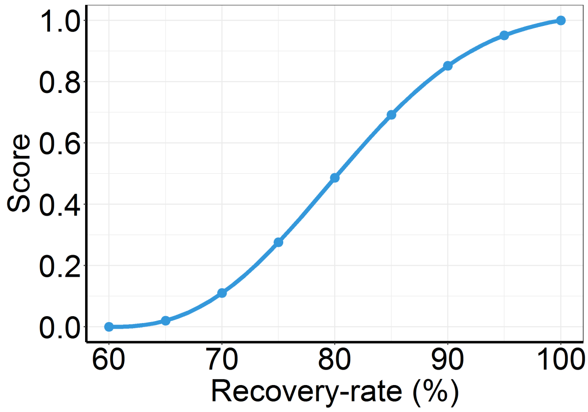


Figure 10.30: SL-Political party's value function for *competence*

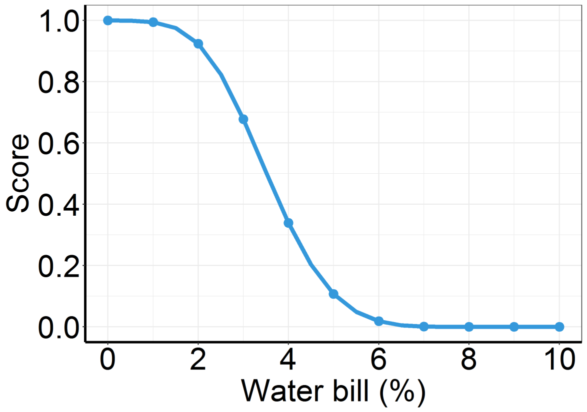


Figure 10.31: SL-Political party's value function for *social justice*

Political demands. Its demands are intended to promote *competence* by means of economic instruments (e.g. increasing water fees) and to ensure *social justice* through economic instruments (e.g. reducing water fees), financial instruments (i.e. social aid), and normative instruments (e.g. water shut-off ban) (Table 10.22)

Table 10.22: Set of political demands of the political party “SL”

Demands	**Preconditions**	**Additional (adv.)**	**Additional (enact.)**	**Framing**	**Type**
Increase tariffs	$C_2 < 0.5$	—	—	$F_2 = 1$	Advocated
Reduce tariffs	$C_4 < 0.5$	$\gamma > 0$	—	$F_4 = 1$	Advocated
Enable social aid	$C_4 < 0.7$	Social aid is disabled.	—	$F_4 = 1$	Advocated

10.8.6 Political party “S”

Values. Its core values are *control* and *social justice* (Table 10.23). This stakeholder considers that *social justice* is extremely important, but unreachable unless there is an effective institutional *control* over the water public service.

Table 10.23: Values of the political party “S”

Value items (ranked)	Weight	Main indicators
Control	3	Public participation in the water utility (%)
Social justice	2	Households with their supply shut off (%); Average share of the income that is used to cover the water bill of those households whose income is lower than 13,000 eur (%)

Value functions:

- **Control**. This function evaluates the world-state using the share of public participation in the water utility (in %) (Fig. 10.34).
- **Social justice.** This *policy-influencer* assesses the world-state with respect to *social justice* using two different value functions. If supply shutoffs are enabled, *social justice* is assessed by means of the number of households whose supply has been shut off (in %) (Fig. 10.32). If supply shutoffs are disabled, *social justice* is assessed by means of the average water bill relative to the income, for those households whose income is lower than 13,000 euros per year (in %) (Fig. 10.33).

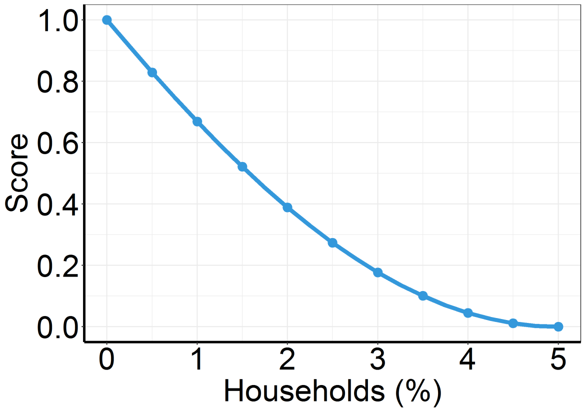


Figure 10.32: S-Political party’s value function for *social justice* if water-supply shutoffs are enabled

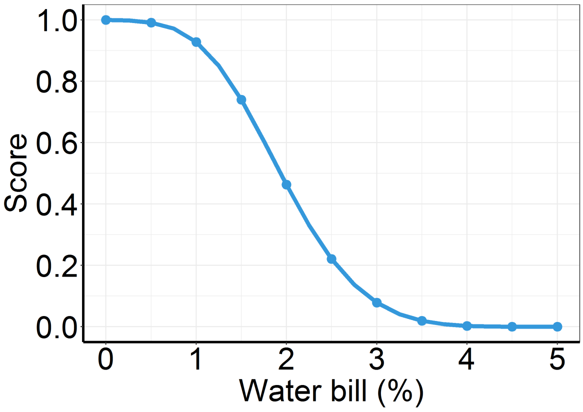


Figure 10.33: S-Political party’s value function for *social justice* if water-supply shutoffs are disabled

Political demands. Its demands are intended to achieve *control* over the water utility (by means of increasing public participation) and to ensure *social justice* through economic instruments (e.g. reducing water fees), normative instruments (e.g. water shutoff ban) and persuading instruments (e.g. social mobilisation) (Table 10.24)

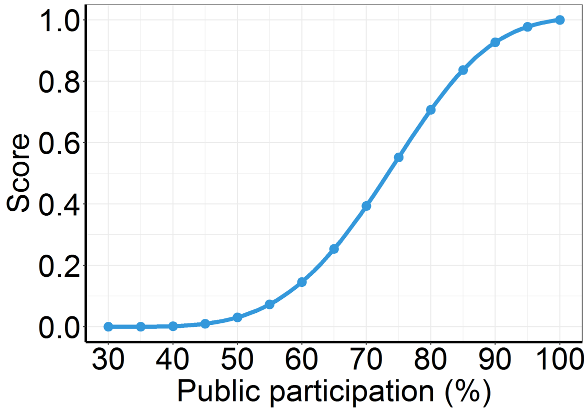


Figure 10.34: S-Political party's value function for *control*

Table 10.24: Set of political demands of the political party "S"

Demands	**Preconditions**	**Additional (adv.)**	**Additional (enact.)**	**Framing**	**Type**
Direct management	$C_1 < 0.5$	Public-participation < 100 %	—	$F_1 = F_4 = F_5 = 1$	Advocated
Disable supply shut off	Always	Supply shut off is enabled.	—	$F_1 = F_4 = 1$	Advocated
Social mobilisation	$C_4 < 0.3$	Mobilised households < 30 %	Social-support > 50 %	$F_1 = F_4 = F_5 = 1$	Advocated + Enactable
End mobilisation	$C_4 > 0.5$	Social mobilisation is active	—	—	Enactable
Reduce tariffs	$C_4 < 0.7$	$\gamma > 0$	—	$F_1 = F4 = 1$	Advocated

10.9 Appendix B: Demands/Actions and effects

Table 10.25: Enactable demands and effects

Demand	Effects
Reduce tariffs	- Decrease factor γ (which multiplies water fees) by 0.10
Increase tariffs	- Increase factor γ (which multiplies water fees) by 0.10
Disable social aid	- Households whose income is below 13,000 are not afforded to benefit from lower water fees.
Enable social aid	- Households whose income is below 13,000 are afforded to benefit from lower water fees.
Disable maximum supply	- Households are afforded to use as much water as they request.
Enable maximum supply	- Households are not afforded to use as much water as they request (namely, they can only use 120 L/p·d).
Disable supply shut-off	- Water Utility is not afforded to shut households' supply off when they are overdue.
Enable supply shut-off	- Water Utility is afforded to shut households' supply off when they are overdue.
Environmental awareness campaign	- *Universalist* households will try to adopt a new *conservation practice*. - *Conservative* households will try to adopt a new *conservation practice* with a 50 % chance. - *Self-centred* households will not adopt any *conservation practice*.
Social mobilisation	- *Self-centred* and *universalist* households that supported *social mobilisation* will not pay their water bills (their debt is not increased).
End mobilisation	- Households that supported *social mobilisation* and did not pay their water bills return to their regular state.
Direct management	- Private participation in the water utility decreases 25 %.
Increase private participation	- Private participation in the water utility increases 25 %.

Part IV

Further work and conclusions

Chapter 11

Further work

The main objective of this dissertation was to understand the role that values play in policy-making and how agent-based simulation can be used to support assessment of value-driven policies. Although the work was driven by scientific and academic values and is so reflected in the body of this dissertation, there was also the empirical motivation of developing technology for artifacts that are useful in actual practice.

This chapter reflects on how what has been accomplished so far can be further developed. Namely, the conceptual contributions of Part II in terms of potential developments for Artificial Intelligence: (i) The use of the framework as a testbed for the "value-alignment problem"; (ii) the design of value-imbued socio-cognitive technical systems that involve hybrid populations of agents (human and artificial); and (iii) two potential venues for the use of the framework (value alignment of norms and value-based argumentation in policy design).

11.1 Value-alignment problem[1]

The popularity of some recent applications of AI has given rise to some concerns in society about the risks of autonomous AI-systems [85, 21, 249]. In response, many scholars and philosophers have focused on the necessity to orientate scientific and technological efforts towards a responsible development of AI, putting "ethics in AI" in the research agenda (e.g. "AI4People" [86], "value-sensitive" [196], etc.).

The conceptual framework can contribute to this end, since it provides a useful double-perspective explore values in AI-backed systems. In particular, by considering values as individual standards that motivate the conduct of the agents and as social standards that concern the society as a whole, it emphasises the two focci of design: (i) one concerns the design of the coordination system (e.g. norms, protocols, etc.) and (ii) the other focuses on the design of behaviour models of agents.

One focus of AI concerns is on the risks associated with the autonomy of artificial entities and the so-called Value-Alignment Problem (VAP), which declares that autonomous entities should behave according to human values [211]. Similar cases have been explored in behavioural ethics in organisations (see [255]) (for instance, ethical leadership, moral entrepreneurship, and ethical culture in businesses [132, 42, 169]). Since AI-backed systems exhibit increasingly complex behaviour, whose conduct have a significant impact on human systems, different strategies may be adopted:

---

[1] A preliminary version of this section was presented in the 18th Mexican International Conference on Artificial Intelligence (MICAI19) (see [186]).

- Design the system so that autonomy is compatible with human values and value-aligned behaviour is fostered (e.g. value-sensitive design [195, 196]);
- Make agents capable of reasoning about moral values and teach them to "do the right thing" (see [21, 15]).

Interestingly, any of these strategies requires a way to "guarantee" that agents behave morally or that autonomous systems are "provably aligned" with a set of values [211]. This constitutes a current important theoretical problem in AI. Although it is unlikely that the value alignment problem may be solved for every autonomous system, it is plausible to develop formal and methodological means to eventually support their provability.

In this spirit, a fundamental contribution of the framework with regard to this problem is its double perspective of values. Supported on the notion of socio-cognitive technical systems —which are constituted of two first-class entities: autonomous agents and a social space—, the framework considers values (i) at a social scale, which result from governing the collective activity of agents; and (ii) at the agent scale, which underlie their reasoning and motivate their individual behaviour in the social space. In other words, agents interact within a social space that is designed so that it ensures that social values are preserved (by means of instruments that "foster" a value-aligned behaviour in agents). In fact, according to this view, policy-making is one particular case of the value-alignment problem: policy design involves steering agent behaviour (which is driven by values), by promoting and discouraging certain actions using instruments such as taxes, subsidies, bans, and information campaigns [75], in order to improve the state of affairs with respect to a constellation of social values. The framework proposes a methodology that is helpful to design systems to "prove" that social values are preserved. By expressing values in operational terms (that is, instantiating them into factual indicators), it is possible to establish value assessment frameworks (basically, design evaluative mechanisms to assess the state of the system with regard to social values).

In this dissertation two case studies in the water domain have been explored using the framework, but there are other paradigmatic problems in AI that can also benefit from this approach. For instance, there is the case of Autonomous Self-driving Vehicles sharing public road infrastructure. In this example, values are reflected (i) in the way autonomous self-driving vehicles are designed to behave (e.g. selfish, cautious, etc), as well as (ii) in the standards that are used to assess the "goodness" of the social outcome (e.g. safety, environmental impact, punctuality, etc). Model-problems include, to name a few, to design decision-making models of autonomous vehicles concerning different values (or hierarchies), and also to choose alternative social values for road-infrastructure governance and their operational definition (for example, "safety" may be evaluated as a function of four factual indicators: number of accidents, traffic density, vehicle-type distribution, and speed limits).

This leads to several research questions: *what values are convenient for the different domains? What variables need to be observable in order to assess the state of the system with regard to a given value? What alternative models can be designed to evaluate and aggregate values? What are the connections between values, goals, motivations, and personality? How can one measure the significance of values, norms, rhetorical messages with respect to a value?*

11.1.1 Design of artificial socio-technical systems

Primiero [200] distinguishes between two type of systems that can simulated by means of ABM:

- **Natural systems.** In these cases, the model aims at mimicking the behaviour of natural agents in naturally occurring environments (i.e. socio-ecological systems). For instance, agents that emulate the human behaviour in an agro-ecosystem, whose social-level outcome —altered by different policies— is the subject of assessment. This approach is used mainly in economics, sociology, ecology, biology and environmental sciences.

- **Artificial systems.** In these cases, the model aims at creating novel artificial environments, where the (social) agents interact. These are more commonly referred to as multi-agent systems (MAS), which are covered in computer science and engineering. This approach is used basically in computer engineering for designing socio-cognitive technical systems for robotics and online platforms (e.g. norms, protocols, etc.).

On the one hand, and following this distinction, one can see the type of agent-based simulation for supporting policy-making in *natural systems*, where policy decisions take values and their consequences into account, by testing and assessing their effects on virtual societies (as illustrated in Chapters 9 and 10).

On the other hand, some AI core ethical concerns can be treated as comprehensive design problems by assuming (i) that autonomous artificial entities are embedded in ecosystems with other artificial and human agents (where social values are essential); (ii) that collective actions involve trade-offs that reflect individuals' objectives and the constraints and affordances defined by of the ecosystem; and (iii) that those trade-offs reflect the values of stakeholders and the ecosystem as a whole.

Thus, (value-driven) agent-based modelling can be useful to frame the problem of *imbuing* values in autonomous artificial entities, from a dual perspective: *imbuing* values in the social coordination system, and *imbuing* values in individuals' behaviour). Simulation offers a "sandbox" for designing *artificial systems* that allows to test and assess the social effects of artifacts, either digital (e.g. social platforms, profiling by means of algorithms, etc.) or hybrid (e.g. smart transportation infrastructure, water technologies, etc.), before implementing such "solutions" in the real world. The key point is the shift of value conflicts approach from isolated agents to socio-cognitive technical systems (multi-agent system). Hence, the focus of design is twofold: (i) modelling a coordination system (e.g. norms, protocols, etc.) and (ii) modelling the behaviour of the agents.

Designers who want (or are requested) to imbue a value into a socio-cognitive technical system can follow the methodology presented in Chapter 8: (1) take those social values to be imbued; (2) contextualise them into the domain and translate them into factual indicators; and (3) explore instruments that lead to improve or preserve assessment-frameworks based on those indicators.

Hence, by considering the conceptual framework and suggestions presented in this dissertation, the design process goes beyond a value-sensitive design [195, 196] and introduces values in the virtual laboratory to test the effects of artifacts in an artificial society. Simulation could identify that some artifacts should be combined with some social coordination instruments —before deploying or releasing the artifact in the real world. Furthermore, this could be supported by online simulators where different stakeholders could test their concerns and question decisions (and also the very artificial society), thereby implementing a space of cooperative design.

11.1.2 Impact assessments of SCTS

Nowadays, there are increasing concerns about the social effects of digital platforms and software solutions. This takes special relevance considering that many of these systems have been implemented without previous assessment of their social effects. While in some engineering fields these concerns have been formalised as professional standards or impact assessment procedures (e.g. *Environmental Impact Assessment* in civil engineering), regulation is being quite slow in the case of digital artifacts.

Digital platforms have a purpose for which the system is designed, but their use may entail undesired social effects (for instance, YouTube provides a platform to watch and share media, but its recommendation algorithm is known to have been abused to recommend increasingly extreme and controversial videos). For this reason, much research aims at taking social values into account to assess the performance of these systems. Interestingly, as Wagner [265] pointed out, the development and adoption of ethical guidelines may be fuelled in order to escape from government regulation, and thus constitute a voluntary non-binding code of conduct (see also [264]).

With this in mind, an *impact assessment protocol* could be a *regulatory requirement* for AI-backed systems to be implemented, leading to similar protocols as the *Environmental Impact Assessment*[2] for engineering projects, involving, in rough terms, the following stages:

1. The promoter presents to the corresponding Authority the initial SCTS-project, including a *request* to determine the scale of assessment.
2. The Authority consults relevant Agencies and stakeholders for suggestions, and afterwards sends a “scope document” to the promoter —which would help the promoter to elaborate the impact assessment document.
3. The promoter sends the SCTS-project and the initial impact assessment document to the Authority.
4. The project is disclosed. Stakeholders can suggest, appeal, or advocate for changes. The projects of this process is then sent back to the promoter.
5. The promoter has to send back to the Authority: (1) a *request* for the project to be assessed; (2) a technical report of the project; (3) an final impact assessment document; and (4) appeals resulting of the project disclosure process.
6. If this application is admitted, the corresponding Agencies will conduct an assessment of the project. The result of this process is an *Impact Assessment Declaration*.
7. If the *declaration* is positive, it has the nature of a prescriptive report, defining under what circumstances the project can be executed, and what are the required preventive, corrective, and compensatory measures.

Generally, a third-party entity carries out the impact assessment, in order to avoid possible conflict of interests. This consultancy service may be explored as a “new” business activity in the next future.

11.2 Value alignment in norms[3]

The conceptual framework of this dissertation provides a “language” and a set of assumptions that delimit the problem in a simple way. Based on this framework, some intuitive applications can be identified and further formalised to provide more abstract problems.

Informally speaking, the Value Alignment Problem can be viewed as to what extent are agents’ decisions (and hence actions) aligned with the values to be imbued. One can use the conceptual framework of the dissertation in the context of norms. In Normative Multiagent Systems, since norms govern behaviour (i.e. decisions and actions) the alignment can be expressed as “suitability” between the norms that govern behaviour and the values that are held. This constitutes an interesting problem: to choose and apply norms in order to achieve valued states of the world.

In this problem, as discussed in [234], the world is modelled as a *Labelled Transition System* (LTS) [95] and is described as a set of states and actions that allow to move from one state to another. As established in the framework, values serve to assess the “goodness” of a state of the world and to decide whether an action is preferable to another. Because of consequentialism and commensurability it is possible to make comparisons and establish preferences between states of the world, and also assess whether the effect on an action produces a “better” state or not —since the state of a SCTS changes

---

[2] `https://en.wikipedia.org/wiki/Environmental_impact_assessment`

[3] A formal approach of this section was presented in the Responsible Artificial Intelligence Agents workshop (RAIA) in AAMAS19 (see [234]). It focused on the alignment of a given system with respect to a given value, for which it explored the relationships between values, actions and norms.

when an agent executes that action. When applying a norm to a given world, it alters the transitions and their resulting states (because certain actions are banned or obliged).

Assuming multiple agents and multiple values, one has to deal with aggregation of value-based preferences from different perspectives. (i) An individual agent may want to aggregate preferences over states with respect to multiple values (i.e. one agent's preference with respect to a set of values). (ii) Similarly, the preferences over the states with regard to one value can be aggregated from a social perspective (i.e. a group of agents' preference with respect to a given value). (iii) Finally, the preferences with respect to multiple values can be aggregated socially (i.e. a group of agents' preference with respect to a set of values).

11.3 Values in argumentation for policy-making[4]

The policy-making cycle involves multiple stakeholders that compete to steer policies towards their (conflicting) interests. For instance, in the domain of agricultural water use, farmers may want a reliable non-restricted supply to ensure the productivity of their farms, while environmentalist organisations want water use to be regulated in order to protect the ecological status of water bodies.

Argumentation is a device that is often used to address these conflicts. Policy-makers and stakeholders argue about the advantages of choosing specific means or to compensate, refine or discard them. Stakeholders need to argue in favour or against specific actions or policy means that may lead to particular outcomes in order to agree on a policy. Interestingly, these arguments in practice tend to involve values, since these are often invoked to justify proposals. As mentioned in previous chapters, values are directly involved in determining whether a state of affairs is “better” and in determining what might be the “right” action to take.

Thus, exploring the dialectical argumentation in value-driven policy-making systems is an interesting research line, particularly useful in the design and negotiation stages of the policy cycle. In this line, the conceptual framework provides a “language” that covers multiple relevant elements, as well as a set of simplifying assumptions. For instance, it assumes a given set of social values, which can be objectively assessed by evaluating some indicators, upon which policy decisions are made. This applies to several situations that involves argumentation: about the operational definition facts and values of a problem, about whether values are satisfactorily met at any point in time or not, or about accepting the results of a simulation, to name a few. As illustrated in [252], the conceptual framework assumes a value-based agent specification that allows to reason with values, allowing to identify what actions can be taken by the agent to reach a target state (deemed desirable with regard to the agent's values).

---

[4] A formal approach of this section was published in Annals of Mathematics and Artificial Intelligence (see [252]). It proposed a formal framework to support the choice of actions of a value-driven agent and arrange them into plans that reflect the agent's preferences.

Chapter 12

Application to water management

Artificial intelligence in general, and agent-based simulation in particular, can provide tools to simulate the effects in controlled (artificial) environments without committing resources and exposing the system to real effects. This approach can be useful to generate new tools in the water industry, as suggested in this chapter —apart from examining socio-hydrological phenomena as shown in previous chapters.

12.1 Agent-based simulation and digital twins

Simulation is increasingly used in the water-industry through the so-called *digital twins* (see [35]). Although they do not use the same approach as agent-based simulation, they do have a similar conceptual approach, since they are based on having a virtual replica of the system where to conduct simulations without disturbing the real system. Technically speaking, a *digital twin* is a platform where a virtual replica of physical assets or processes involved in the life cycle of a product or service is built to optimise its management. However, the focus of *digital twins* is the long-term control and management of physical assets through predictive maintenance. Thanks to the Internet of Things (IoT), physical assets are equipped with devices that can capture a variety of data in real-time, constantly updating the *digital twin* (which acts also as a repository of data). These huge databases are then treated using *machine learning* (ML) to predict possible failures and breakdowns, thus avoiding setbacks and production losses by anticipating such events. In the water sector, *digital twins* are being developed for management of water infrastructure (e.g. wastewater treatment processes, supply networks, etc.).

While *digital twins* are focused on the management of the physical infrastructure based on historical data, ABMs can further expand such approach by including socio-cognitive aspects involved in the systems and by considering new scenarios. In ABM, agents' behaviour and networks are explicitly modelled and not computationally inferred by data, and thus it is fully explanatory and can be contrasted with theoretical insights and empirical evidence [92, 96, 261]. As long as the models are sound, ABM are capable to generate 'what-if' scenarios that produce previously-unseen events. Hence, they may be a powerful tool to generate synthetic datasets or explore different management strategies or operating environments. Interestingly, Minsker *et al.* [166] pointed out that the consequences of climate change are invalidating many of the design assumptions about the operating environment in which infrastructure are supposed to operate (for instance, the meteorological regime), making urgent to explore new design and support-tools: thus, agent-based simulation could be used to explore the performance of some designs under diverse scenarios).

Yet, the commercial exploitation of agent-based simulation is not widespread. Only a few companies provide consulting for strategic decision-making by means of new-scenario generation and models of social behaviour (for example, customers' brand selection, prediction of election results, etc).

Based on these insights, and beyond the context of digital twins, the hydraulic design and management of physical assets could introduce social behaviour into account, either considering technical components of industrial processes as social agents (IoT) or reproducing the water-related behaviour of the neighbours of a community under different events. This could lead to develop modules of social behaviour for traditional hydraulic engineering software (e.g. SWMM for sanitary and sewage systems, or EPANET for supply systems).

The inclusion of the socio-cognitive aspects and behaviour of citizens is deemed essential for managing smart cities, because understanding human behaviour and reasoning (in choices, decisions, mobility, activities, etc.) is key in design and management of city services. In fact, forecasting water-demand has been broadly covered in agent-based models [158, 170, 88]. Also, there are European projects that focus on "sewer sociology"[1] in order to improve the knowledge the social habits on waste management from data of the sewage system. Accordingly, agent-based models could support the design and management of the water supply or sanitation networks [32]. In this line, some works have focused on segmenting households' behaviour with respect to water public services [41, 45]. Different behavioural models can be built based on these generic clusters, which could be used in water-domain agent-based models (exploring, for instance, weather singular events or housing future trends).

Interestingly, agent-based models could serve to generate synthetic data at a lower cost (conducting simulations is faster and at less cost than fieldwork)—which is particularly useful to create scenarios with no historical precedent that could serve to train ML-systems [107].

Furthermore, agent-based simulation could serve to assess the scalability of prototypes, as require many funding programmes (e.g. EU Life programme[2]) (see Sec. 11.1.1).

12.2 Value-driven management of water and other services

As Simon [236] pointed out, when conflict and uncertainty are present, professionals, institutions and society should collaborate to design artifacts that aim at achieving those desirable states of affairs. The reason is that members of society are not passive agents: they react to the plans the professionals propose because they seek to use the system to further their own goals. Since the social behaviour may alter the outcome of plans and eventually lead to unintended states, it is an important factor to be taken into account while designing and monitoring management policies. Following a similar view, Shindler *et al.* [233] researched social acceptance of management practices. Even further, Brouwer *et al.* [41] mentioned a mission shift in water companies, from optimising the supply of water (i.e. safe and high-quality water at an affordable price) to (also) adapting the water service to the citizen demands (i.e. improving the service by taking citizens' values, goals and motivations into account).

The conceptual framework in this thesis can support the exploration of cases in connection with this new mission. As presented, values are involved in the definition and resolution of social problems. In simple terms, values play a role in setting ends and its practical implementation through indicators, and in setting the means and instruments that will have a positive effect on the state. Then, identifying what values are embedded in or are to be imbued into water management, and how they can be operationalised, is a relevant issue for companies in the water sector [133]. Having this clearly defined can lead to better-designed support systems that support public decisions and management policies, by better contextualising the values and societal goals pursued and establishing value-diverse

---

[1] Corominas, L. (coordinator): `https://www.scorewater.eu/cases/barcelona-en`

[2] `https://ec.europa.eu/easme/en/life`

indicators [34] —thus contributing to transparency and accountability and, eventually, to stimulate communicative strategies.

In fact, some managers may question the fact that their decisions are based on values, since they think it may undermine the quality of the management —which is, supposedly, objective and apolitical because it is based only on technical knowledge and and "rational economics" (see [44]). However, those decisions are not based only on knowledge of the system, but also on values [237, 185]. Thus, values are present in management decisions, regardless of them being consciously or unconsciously taken into account. In the most simple terms, a decision-maker uses some sort of construct (i.e. values) to discern whether some state is satisfactory or not, and to elicit preferences over states of the world (see [135]). In a more technical view (as discussed in Chapter 8), this consists of choosing a particular operational definition of the problem and the decision-making criteria (i.e. values) —which involves a choice about what is relevant and important. Interestingly, the fact that values are involved in such decisions is not wrong *per se*: values express standards of "desirability" at a collective level (i.e. the "social value") and are worked out to define trade-offs (i.e. hierarchies and relative importance). However, it is important to recognise that there are value conflicts in those management decisions, which clearly can lead to social acceptance issues if the values that guide management are not shared social values [44]. Noteworthy, these issues are embedded in agent-based models that are used as decision-support tools.

For instance, under a traditional economic paradigm, the urban water service may be viewed as service for a customer, emphasising specific interpretations of values such as 'efficiency' and 'excellence'. In this view, innovation has been steered towards reducing fees and industrial costs. However, there may be other values that should be considered in the definition of the problem (and not only when thinking on mitigation measures), such as 'social equity' or 'loyalty'. This way, these values would not be incorporated only as a corporation code of conduct, but also as management ultimate goals. In practice, the computational model could support decision-making so that the management practices are aligned with those social values, which are operationally translated into indicators and thresholds. The management of the ecosystem could be supported by an online platform and sensors that monitor multiple indicators (thus, values). Agents' behaviour or external/internal events may drive the system to "unsatisfying states" (i.e. some indicators have dropped below satisficing levels), in which cases the manager could discuss with the community possible *means* in order to satisfy their *ends*.

In the terms of the conceptual framework, this view considers:

- Social values in the *social space* where multiple agents interact (e.g. watershed, urban water-supply network, urban water-sanitation network, etc.). For example, 'equity', 'innovation', 'safety', 'efficiency', and 'environmental protection'.
- Agent values that participate in the *social system* (e.g. water users, wastewater treatment plants, workers, etc.).

Thus, for instance, a *policy-schema* may declare 'equity' and 'safety' to be imbued in the urban water-supply system under an extreme scarcity scenario. Its *policy ends* define a *state of the world* where users share the water resources in an *equitable* way, using as *indicators* the 'litres per person per day' (mean and distribution), but also considering that people are *safe* by monitoring the 'number of people below a safe supply-level'. Given this purpose, the *policy means* rely on 'managing users' water demand', with *instruments* (e.g. norms and incentives) that sanction users when their water usage exceeds a threshold, blocking the supply over the threshold if this occurs repeatedly. This system could be tested considering different populations of agents (i.e. different value-driven individual behavioural models) to assess their performance. The outcome could help to identify satisfactory or counterproductive management-strategies or necessary interventions in the real-world (for instance, campaigns to change some habits in human users, re-design of autonomous technological devices, etc.).

This approach entails multiple challenges:

- Building a valid model, suitable for simulating the social effects of different water management policies. This involves expert knowledge, field work, data gathering and analysis, exploration of models of human behaviour, etc., but also new departments in the company (e.g. computer scientists, programmers, sociologists, AI experts, etc.) and appropriate technical infrastructure (e.g. databases, cloud simulators, etc.). And, finally, adjusted to new regulation and ethical codes (e.g. privacy, data protection, use of models, etc.).
- Whose and what values (e.g. philosophers, experts, politicians, stakeholders, scientists, etc.), how these are agreed on (e.g. public participation, mediation, etc.), and how they are operationalised.
- What specific strategies are adopted to assess solutions (e.g. short-term or long-term, territorial scale, etc.).
- Appropriate technological tools and algorithms to address these type of problems.
- Human experts who are trained with simulation-based tools in order to provide appropriate consultancy.

Chapter 13

Conclusions

This dissertation's main objective was to outline a conceptual framework for the use of simulation in policy-making. Motivated by the urgency to address threats to human civilisation which require collective action, the presented conceptual framework may contribute to build decision-support systems for management and policy-making.

This chapter enumerates the main insights of the dissertation. These are organised again in three blocks —that correspond to policy-making, Artificial Intelligence and water domain— and, for each block, comments are referred back to the research questions stated in Section 1.4.

13.1 Conclusions for simulation for policy-making

Policy-making is a value-driven social activity. Regardless of the collective problems are solved by means of technologies (say, for instance, smart city solutions) or other instruments, there are values underlying the problem-solving process. Values reflect the non-questioned core beliefs that underlie design and policy decisions (thus, relate to political and cultural biases), but they also indicate the notions of good and worthiness. Accordingly, knowing the role that values play in decisions can lead to refine such concepts, thereby improving policy and design decisions.

Agent-based simulation for policy-making, which is covering increasing attention due to their potential, is subject to values, as well. From the abstraction of the real world system (i.e. representation of the problem) to the assessment of policies and resulting states (i.e. assessment of solutions and effects), values are present. However, little literature addresses the challenges that imply working with values, especially in computational models for supporting decision-making, leaving many modelling choices without background material (e.g. how to measure and infer values, how to model them, how to implement value trade-offs, etc.). This justifies a conceptual framework that enlightens how values are involved in agent-based simulation for policy-making.

These are the main conclusions of this dissertation on the topic of simulation for policy-making:

- Collective problems require governance in order to coordinate social systems constituted of large numbers of agents. Since the behaviour of societal systems cannot be fully predicted or controlled due to their complexity, simulation of policies may support the design of public policies because it has potential in informing policy decisions by anticipating situations and increasing knowledge about social phenomena. However, one must understand that simulation is conducted on a model that defines the policy problem, especially who is going to be incentivised (i.e. *policy-targets*), for what (i.e. *policy-ends*), and how (i.e. *policy-means*). (**Q1**)

- Values are standards that express the worthiness, the desirability, and the goodness of something. In practice, values are cognitive constructs that are involved in supporting moral intuitions and ethical beliefs; in perceiving and evaluating objects and environment; in deriving individuals' core preferences; and in interacting with other stakeholders. Thus, values have much potential for sociological research. Besides, they are assumed to be an essential entity in policy-making: values justify interventions that lead to particular state of affairs of the societal system —which may lead to disputes in a political arenas (e.g. whether some values should be pursued, how they should be understood, whether some values are more important than others, etc.)— and values are involved in agents' behaviour and acceptance. (**Q1**)

- Cognitively, values activate an intuitive association of abstract ideas to expected outcomes. However, values do not have an univocal meaning. Because of this association, some aspects of the environment are perceived, assessed, and eventually decided over. Interestingly, the fact that values do not have an unambiguous meaning do not make them useless; rather, they are useful because activate moral intuitions in complex decision situations that cannot fully grasped by factual information, as are policy decisions. In fact, since values are invoked to justify policy and management decisions, they should be taken into account in simulation for policy-making. (**Q1**)

- Three major entities are important to understand value-driven policy-making: (i) the individuals, (ii) the society (supra-individual entity), and the (iii) policy that coordinates the behaviour of individuals towards some social state of affairs. For policy design, this perspective conceives policy instruments as mechanisms that relate to individuals' *motivational values*, understanding that their behaviour —driven by those values, as explains the Schwartz's theory of values— has an effect onto the *social values* —that express aspects of the common good and public interest. This shows the interdependence between (the values) of individuals and society. (**Q1**)

- Values do not give an exact information about policies and state of affairs on their own. Although it is true that abstract values incorporate certain characteristic expectations, these are not operational for policy assessment with computational models —unless a proper contextualisation and translation process is done, as suggests this dissertation (see next section). In essence, value judgements are subject to particular value understandings, and therefore they should be made explicit and transparent in decision support systems as much as possible. Hence, values cannot remain in an abstract state in artifacts: they have to be instantiated in observable terms that make possible to prove that some action or intervention has an effect on those terms. Furthermore, in policy-making, since particular social states are promoted over others based on values, these values should projected onto observable indicators (qualitative or quantitative). (**Q1, Q2, Q3, Q4**)

- The conceptual framework presents the main elements to design a "policy problem", which are the *domain* of intervention, the *stakeholders*, the *values*, and the *policy-schema*. The *policy-schema* is a simple way to model policy interventions, distinguishing between two main components: (i) *policy ends* (i.e. *intended effect*), which are those objectives that are meant to be achieved with the intervention (i.e. those *better* states of affairs); and (ii) *policy means* (i.e. *potential cause*), that are those instruments that, one implemented in the social space, will produce a change on the state of affairs. (**Q1**)

- The conceptual framework proposes to assess policies (i.e. *policy-schemas*) with respect to *effectiveness*, *desirability*, and *acceptability*. Policies are *effective* when its actual outcome is consistent with its intended outcome; are *desirable* with regard to the values of a stakeholder (namely when the actual outcome is deemed "good"); and are *acceptable* when stakeholders agree that the outcome is sufficiently similar to the outcome of the most-known favourable alternative intervention (including non-acting). Interestingly, these notions are judgements that are supported on other value judgements (implying particular definitions of other values). (**Q4**)

— Accordingly, an essential point of governance is to establish how social values are going to be understood (in order to be preserved and improved). Without a proper system to monitor these values, they are at risk of being ignored, which is specially true for computational models and digital platforms. In other words, if values cannot be properly elicited or observed, they may be neglected. Consequently, setting a particular operational definition for values and its corresponding evaluation methods involves to make a decision about what is relevant and desirable, but also clarifying the stances on design, management, and policy-making. Therefore, there are two tasks when addressing policy-making with computational models: (i) the designation of the values to be promoted and preserved; and (ii) the definition of these values and (iii) eventual translation into indicators. However, collective problems are, generally, wicked problems, which means that there is no objective assessment framework nor a perfect solution to the problem. This indicates that stakeholders should be flexible (e.g. novel interpretation of problems) and open to cooperate with each other (e.g. share knowledge). (**Q1, Q2, Q3, Q5**)

13.2 Conclusions for Artificial Intelligence and Multi-agent systems

This dissertation has presented a way to bring values into artificial intelligence, namely into multi-agent systems modelling, by taking insights from different research fields. In simple terms, values are taken as a construct that is involved in the perception, evaluation (i.e. standards or criteria) and decision-making. This has implications for the agents and the behaviour they exhibit, and also for the coordination of the multi-agent system as a whole.

Thus, although values are not the only cognitive construct involved in decision-making, they are deemed essential when making major decisions, especially those that concern multiple agents —recognising some sort of social welfare. An artificial agent, although it does not have a concrete moral system, seeks to achieve particular states of affairs (which are deemed desirable by its designer) in a social world constituted of other artificial entities and human beings. Likewise, those in charge of coordinating the system make decisions in accordance with their values, since they promote certain states of affairs over others. In the same way that a designer "imbues" an artificial agent with its values, the social system may be imbued with values if its policies lead to value-aligned outcomes (proven by assessment frameworks).

These are the main conclusions of this dissertation for artificial intelligence:

— In multi-agent systems, values can be approached from a double perspective: (i) first, values that result from governing the activity of agents at the social scale (i.e. *social values*); (ii) second, values that guide the individual behaviour of agents in the society, which are included in their socio-cognitive models (i.e. *motivational values*). Policies coordinate agents' behaviour towards some *better* social state of affairs, thus "bridging" the interaction between *motivational values* and *social values*, and hence between the heterogeneous *agents* and the *society*. (**Q1**)

— Values are required to be instantiated into operational constructs, meaningful for the domain and context of interest. In the present dissertation, a suggested practical way to work with values in computational models for policy-making has been to translate abstract values into observable indicators and assessment frameworks (i.e. evaluation mechanisms and aggregation mechanisms). Accordingly, coordination policies, in order to imbue the system with values, must be able to "improve" those variables that characterise the state of the system and are associated with the values to be imbued. However, the actual state of the system may be not perceived by agents as the manager (or policy-maker) of the system does. That is, agents' perception may be partial (not all the variables of the world-state are relevant or observable); focused on a particular scale (either local or global); subjective (different indicators are used to assess objects with regard to the values they hold, that the manager may not know); and complemented with other knowledge besides value-associated indicators. (**Q2, Q3, Q5**)

— The methodology of (i) taking abstract values; (ii) contextualise them in the domain; (iii) translate them into factual indicators; and (iv) define the (coordination) problem (i.e. means and ends) around this, may be a useful approach for *imbuing* values to hybrid systems, which is a current problem in AI nowadays. (**Q2, Q5**)

— Problems of designing coordination policies for multi-agent systems may be framed as *design problems* [236,185], which are constituted of (i) means (i.e. command variables); (ii) ends (i.e. standards of assessment and constrains); and (iii) laws (i.e. causal mechanisms and parameters). These problems can be addressed computationally in order to search for optimal (or satisficing) solutions. Interestingly, defining such problems involves, necessarily, values (for instance, by establishing the standards by which solution proposals are going to be assessed with), especially in the "representation problem", which has to do with the ontology that defines the world and that enables to assess the solutions with regard to different values (because there exist variables onto which it is possible to project those values). (**Q1**)

— Recognising the role of values in approaching coordination problems is relevant for one plain reason. It identifies that behind *optimising* an *objective function* there are values —thus, one should be careful when addressing real-world social problems. Taking a step back may be useful to recognise that other values are involved in the problem and therefore should be included in the *objective function* to *optimise*. (**Q5**)

13.3 Conclusions for the water domain

Following the outlined conceptual framework, two study cases in the water domain have been explored.

The first case study explores a community of farmers who are to adopt an irrigation technology. This problem is usual in Spanish water-agriculture public policies, which promote the irrigation agriculture to achieve, eventually, the socio-economic development of rural communities to curb depopulation. With this in mind, the simulation model focuses on one specific intervention aiming at imbuing "modernity" to the farming community. In this spirit, the diffusion and adoption of high-tech irrigation systems is seen as desirable (i.e. policy end) and is promoted by means of economic incentives (i.e policy means), because farmers are viewed as profit-seeking agents because they are part of agricultural systems with a strong presence of market-oriented agro-industries.

The second case study explores the governance of the urban water public service. In it, multiple stakeholders with different values and political beliefs evaluate the state of the world and try to produce policy shifts so as to contribute to their agendas. Furthermore, households act in that space to use the water supply service for their needs, but also participating in the political arena, since they also evaluate the world-state and support political demands. In this case, the model focuses on the simulation of a social behaviour as a result of the existence of competing value-driven stakeholders in the urban water domain.

These are the main conclusions of this dissertation for the water domain:

- The contribution of the conceptual framework presented in this dissertation lies in emphasising the existence of (i) the values of stakeholders (i.e. what they consider relevant and base their decisions on, how they understand such values in a socio-hydrological context, what actions do they motivate, etc.); and (ii) the social values, to be promoted by policy interventions (i.e. what values are to be imbued in the societal system and why, how are they expressed and understood, how are they measured, what instruments are to be implemented for that matter, etc.); which are both important for addressing collective problems in the water domain. In fact, a recent research line in water studies aims at exploring a segmentation of water customers based on "soft" differences (e.g. lifestyles, personal views, values, norms, goals, motivations), which can be complementary to those classical segmentation based on "hard" socio-demographic differences [41]. This dissertation provides an exploration of values in policy problems and suggests how to address them with agent-based simulation. (**Q1, Q6**)

- The double-approach of values (i.e. *social values* and *individual values*) is particularly useful in the water domain, because it emphasises the twofold view of water governance: at the agent level (e.g. users, citizens, customers, etc.), but also at the context in which the agent operates —space that is shared with other entities, and therefore requires social coordination— (e.g. watershed, community, user association, etc.). This view can be useful to address water crisis and manage water services. (**Q1, Q6**)

- The social outcomes are output of a complex system (e.g. multiple stakeholders, multiple stances, many factors, etc.). Due to the complexity and the significance of certain social problems, simulation is a promising tool for decision support systems. Since management of water public services has to consider multiple values and competing interests, simulation may be useful to test decisions in a controlled environment, before implementing such decisions in the real world. This can help to avoid counterproductive situations and social acceptance issues. Presumably, such management support systems not only require technical infrastructure (e.g. online simulators, servers, reporting tools, etc.), but also constant fieldwork (e.g. sociological research, participatory workshops, etc.) and field experts (e.g. sociologists, engineers, etc.). In this connection, two case studies in socio-hydrology have been addressed with agent-based simulation, showing the potential use for the water domain. (**Q6**)

- Managers of public services often reject the fact that their decisions are based on values, as they think it may undermine the quality of the management —which is, supposedly, objective and apolitical because it is "technical" and based on "sound science" and "rational economics". Nonetheless, values are present in the frameworks that underlie management decisions, regardless of them being consciously or unconsciously taken into account. Choosing any particular operational definition of decision-making criteria (that is, values) involves making a choice about what is relevant and important. This is why any decision concerning a societal system (as involves the management of water services) is characterised by conflicts between competing values, perspectives, interests and those stakeholders that represent these stances. As a consequence, deemed-optimal management is perceived as such within a particular definition of the management problem, which may be questioned by other stakeholders. And, clearly, management practices that lack social acceptance are vulnerable to public rejection. Although many management decisions are based on central values like efficiency, other values may be relevant to ensure the (social, environmental, and economic) sustainability of water public services, since citizens' preferences may be derived from social values that are not traditionally considered in management decisions. (**Q1, Q3, Q6**)

- As a consequence, management problems may suffer from a myopic view. Many managers think that problems and acceptance issues are because of a lack of efficiency, but the truth is that they arise because other values are relevant for the stakeholders in the domain (values that may not fit the economy-reductionist perspective). Thus, political actors (i.e. *policy-influencers*) participate in political arenas to question the decision-making frameworks and to advocate for including their values in the management policies. (**Q1, Q6**)

References

[1] Aaron, H.J.: Distinguished lecture on economics in government: Public policy, values, and consciousness. Journal of Economic perspectives 8(2), 3–21 (1994)

[2] Aarts, H.: Goals, motivated social cognition and behavior. In: Fiske, S.T., Macrae, C.N. (eds.) The SAGE Handbook of social cognition, pp. 75–79. SAGE Publications Ltd (2012)

[3] Achen, C., Bartels, L.: Democracy for realists: Why elections do not produce responsive government. Princeton Studies in Political Behavior, Princeton University Press (2016)

[4] Agència Catalana de l'Aigua: Canon social del agua *(in Spanish)*, [Online. Retrieved 2019, January 16] `http://web.gencat.cat/es/tramits/tramits-temes/Canon-social-de-laigua-00001?category=75b0b894-a82c-11e3-a972-000c29052e2c&evolutiuTramit=1`

[5] Agència Catalana de l'Aigua: Pla de gestió del districte de conca fluvial de Catalunya i Programa de mesures. 2016-2021. Document de síntesi (*in Catalan*). Tech. Rep. B 10794-2017, Generalitat de Catalunya (2017)

[6] Ahrweiler, P.: Agent-based simulation for science, technology, and innovation policy. Scientometrics 110(1), 391–415 (2017)

[7] Aigües de Barcelona: Precios y Tarifas *(in Spanish)*, [Online. Retrieved 2019, January 16] `http://www.aiguesdebarcelona.cat/facturadelaigua/es/precios-tarifas/`

[8] Akhmouch, A., Correia, F.N.: The 12 OECD principles on water governance — When science meets policy. Utilities Policy 43, 14–20 (2016)

[9] Alarcon, B., Aguado, A., Manga, R., Josa, A.: A value function for assessing sustainability: Application to industrial buildings. Sustainability 3(1), 35–50 (2011)

[10] Alcón, F.: Adopción y difusión de las tecnologías de riego: aplicación en la agricultura de la Región de Murcia *(in Spanish)*. Ph.D. thesis, Universidad Politécnica de Cartagena (UPCT), Spain (2007)

[11] Alcón, F., Tapsuwan, S., Brouwer, R., de Miguel, M.D.: Adoption of irrigation water policies to guarantee water supply: A choice experiment. Environmental Science & Policy 44, 226–236 (2014)

[12] Aldewereld, H., Boissier, O., Dignum, V., Noriega, P., Padget, J.: Social Coordination Frameworks for Social Technical Systems. No. 30 in Law, Governance and Technology Series, Springer International Publishing (2016)

[13] Alhama de Murcia IC: The Tajo-Segura transfer and irrigation (in Spanish) (2016), [Online. Retrieved 2017, August 21] `https://www.cralhama.org/el-trasvase-tajo-segura-y-los-regadios/`

[14] Alhama de Murcia IC: Distribution of crops, tenure and irrigation systems (in Spanish) (2018), [Online. Retrieved 2019, May 28] `https://www.cralhama.org/distribucion-de-cultivos-propiedad-y-sistema-de-riego/`

[15] Allen, C., Varner, G., Zinser, J.: Prolegomena to any future artificial moral agent. Journal of Experimental & Theoretical Artificial Intelligence 12, 251–261 (2000)

[16] Allen, R.G., Pereira, L.S., Raes, D., Smith, M.: Crop evapotranspiration — Guidelines for computing crop water requirements. United Nations (1998), FAO Paper 56

[17] An, L.: Modeling human decisions in coupled human and natural systems: Review of agent-based models. Ecological Modelling 229, 25–36 (2012)

[18] Andrighetto, G., Governatori, G., Noriega, P., van der Torre, L.W.N. (eds.): Normative Multi-Agent Systems, vol. 4. Dagstuhl Publishing (2013)

[19] Aodha, L.n., Edmonds, B.: Some pitfalls to beware when applying models to issues of policy relevance. In: Edmonds, B., Meyer, R. (eds.) Simulating Social Complexity: A Handbook, pp. 801–822. Springer International Publishing, Cham (2017)

[20] Atkinson, K., Bench-Capon, T.: Practical reasoning as presumptive argumentation using action based alternating transition systems. Artificial Intelligence 171(10), 855–874 (2007)

[21] Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.F., Rahwan, I.: The moral machine experiment. Nature p. 1 (2018)

[22] Bainbridge, W.S., Brent, E.E., Carley, K.M., Heise, D.R., Macy, M.W., Markovsky, B., Skvoretz, J.: Artificial social intelligence. Annual Review of Sociology 20, 407–436 (1994)

[23] Banks, J.: Handbook of Simulation. John Wiley & Sons, Inc. (1998)

[24] Bannister, F., Connolly, R.: ICT, public values and transformative government: A framework and programme for research. Government Information Quarterly 31(1), 119 – 128 (2014)

[25] Bardi, A., Schwartz, S.H.: Values and behavior: Strength and structure of relations. Personality and Social Psychology Bulletin 29(10), 1207–1220 (2003)

[26] Baumeister, R.F., Vohs, K.D.: Self-regulation, ego depletion, and motivation. Social and Personality Psychology Compass 1(1), 115–128 (2007)

[27] Becu, N., Perez, P., Walker, A., Barreteau, O., Page, C.L.: Agent based simulation of a small catchment water management in northern Thailand: Description of the CATCHSCAPE model. Ecological Modelling 170(2), 319–331 (2003)

[28] Bench-Capon, T., Modgil, S.: Norms and value based reasoning: Justifying compliance and violation. Artificial Intelligence and Law 25(1), 29–64 (2017)

[29] Benouniche, M., Kuper, M., Hammani, A., Boesveld, H.: Making the user visible: analysing irrigation practices and farmers' logic to explain actual drip irrigation performance. Irrigation Science 32(6), 405–420 (2014)

[30] Beratan, K.K.: A cognition-based view of decision processes in complex social–ecological systems. Ecology and Society 12(1), 27 (2007)

[31] Berger, T.: Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agricultural Economics 25(2), 245–260 (2001)

[32] Berglund, E.Z.: Using agent-based modeling for water resources planning and management. Journal of Water Resources Planning and Management 141(11) (2015)

[33] Boelens, R., Hoogesteger, J., Swyngedouw, E., Vos, J., Wester, P.: Hydrosocial territories: a political ecology perspective. Water International 41(1), 1–14 (2016)

[34] Bond, S.D., Carlson, K.A., Keeney, R.L.: Generating Objectives: Can Decision Makers Articulate What They Want? Management Science 54(1), 56–70 (2008)

[35] Boschert, S., Rosen, R.: Digital twin — the simulation aspect. In: Hehenberger, P., Bradley, D. (eds.) Mechatronic Futures: Challenges and Solutions for Mechatronic Systems and their Designers, pp. 59–74. Springer International Publishing, Cham (2016)

[36] Bostrom, A., Fischhoff, B., Morgan, M.G.: Characterizing mental models of hazardous processes: A methodology and an application to radon. Journal of Social Issues 48(4), 85–100 (1992)

[37] Botterill, L.C., Fenna, A.: Interrogating Public Policy Theory. Edward Elgar Publishing, Cheltenham, UK (2019)

[38] Botterill, L.C., Hindmoor, A.: Turtles all the way down: bounded rationality in an evidence-based age. Policy Studies 33(5), 367–379 (2012)

[39] Bozeman, B.: Public Values and Public Interest: Counterbalancing Economic Individualism. Georgetown University Press (2007)

[40] Bridgman, T., Cummings, S., Ballard, J.: Who Built Maslow's Pyramid? A History of the Creation of Management Studies' Most Famous Symbol and Its Implications for Management Education. Academy of Management Learning & Education 18(1), 81–98 (2019)

[41] Brouwer, S., Pieron, M., Sjerps, R., Etty, T.: Perspectives beyond the meter: a Q-study for modern segmentation of drinking water customers. Water Policy 21(6), 1224–1238 (11 2019)

[42] Brown, M.E., Treviño, L.K., Harrison, D.A.: Ethical leadership: A social learning perspective for construct development and testing. Organizational Behavior and Human Decision Processes 97(2), 117–134 (2005)

[43] Brown, P., Schmidt, J.: Water Ethics: Foundational Readings for Students and Professionals. Island Press (2010)

[44] Brunson, M.W.: A definition of "social acceptability" in ecosystem management. In: Defining Social Acceptability in Ecosystem Management, pp. 7–16. United States Department of Agriculture (1996)

[45] Burns, A.V., Vardy, S., Rowe, M., Murchison, R.: Victorian Water Customer: Water Services Needs & Values Summary Report. Tech. rep., Accenture, Melbourne (2011)

[46] Cai, X., Wallington, K., Shafiee-Jood, M., Marston, L.: Understanding and managing the Food-Energy-Water nexus — Opportunities for water resources research. Advances in Water Resources 111, 259–273 (2018)

[47] Cairney, P.: The Politics of Evidence-Based Policy Making. Palgrave Macmillan UK (2016)

[48] Camill, P.: Global change. Nature Education Knowledge 3(10), 49 (2010)

[49] Campbell, J.L.: Institutional analysis and the role of ideas in political economy. Theory and Society 27(3), 377–409 (1998)

[50] Campbell, J.L.: Ideas, politics, and public policy. Annual Review of Sociology 28(1), 21–38 (2002)

[51] Carrington, D.: England could run short of water within 25 years (2019, March 19), [Online. Retrieved 2019, March 20] `https://www.theguardian.com/environment/2019/mar/18/england-to-run-short-of-water-within-25-years-environment-agency`

[52] Castelfranchi, C.: Simulating with Cognitive Agents: The Importance of Cognitive Emergence. In: Sichman, J.S., Conte, R., Gilbert, N. (eds.) Multi-Agent Systems and Agent-Based Simulation. pp. 26–44. Springer Berlin Heidelberg (1998)

[53] Castelfranchi, C.: Making Visible "the Invisible Hand". The Mission of Social Simulation. In: Adamatti, D.F., Dimuro, G.P., Coelho, H. (eds.) Interdisciplinary Applications of Agent-Based Social Simulation and Modeling, pp. 1–19. IGI Global (2014)

[54] Castro, J.E.: Water governance in the twentieth-first century. Ambiente & Sociedade 10(2), 97–118 (2007)

[55] CEDEX: Evaluación del impacto del cambio climático en los recursos hídricos en régimen natural *(in Spanish)*. Tech. Rep. CEDEX 42-407-1-001, Spanish Government, Madrid (2010)

[56] Chen, X., Lupi, F., An, L., Sheely, R., Viña, A., Liu, J.: Agent-based modeling of the effects of social norms on enrollment in payments for ecosystem services. Ecological Modelling 229, 16–24 (2012)

[57] Committee on Standards in Public Life (United Kingdom Government): Ethical standards for providers of public services (June 2014)

[58] Committee on Standards in Public Life (United Kingdom Government): The Continuing Importance of Ethical Standards for Public Service Providers (May 2018)

[59] Conte, R., Gilbert, N., Bonelli, G., Cioffi-Revilla, C., Deffuant, G., Kertesz, J., Loreto, V., Moat, S., Nadal, J., Sanchez, A., Nowak, A., Flache, A., San Miguel, M., Helbing, D.: Manifesto of computational social science. The European Physical Journal Special Topics 214(1), 325–346 (2012)

[60] Costanza, R.: Social goals and the valuation of ecosystem services. Ecosystems 3(1), 4–10 (2000)

[61] Coward, E.W.: Planning technical and social change in irrigated areas. In: Cernea, M.M. (ed.) Putting people first: sociological variables in rural development, pp. 46–72. 9831, Oxford University Press, Washington, D.C. (1991)

[62] DeCaro, D.A., Arnold, C.A.T., Frimpong Boamah, E., Garmestani, A.S.: Understanding and applying principles of social cognition and decision making in adaptive environmental governance. Ecology and Society 22(1), 33 (2017)

[63] Deffuant, G., Huet, S., Amblard, F.: An individual-based model of innovation diffusion mixing social value and individual benefit. American Journal of Sociology 110(4), 1041–1069 (2005)

[64] Deffuant, G., Huet, S., Bousset, J., Henriot, J., Amon, G., Weisbuch, G., et al.: Agent based simulation of organic farming conversion in allier département. In: Janssen, M.A. (ed.) Complexity and Ecosystem Management: the theory and practice of multi-agent systems, pp. 158–189. Edward Elgar Publishers (2002)

[65] Dente, B.: Understanding Policy Decisions. SpringerBriefs in Applied Sciences and Technology, Springer International Publishing (2013)

[66] Di Baldassarre, G., Sivapalan, M., Rusca, M., Cudennec, C., Garcia, M., Kreibich, H., Konar, M., Mondino, E., Mård, J., Pande, S., Sanderson, M.R., Tian, F., Viglione, A., Wei, J., Wei, Y., Yu, D.J., Srinivasan, V., Blöschl, G.: Sociohydrology: Scientific Challenges in Addressing the Sustainable Development Goals. Water Resources Research 0(55), 0 (2019)

[67] Dietz, T., Fitzgerald, A., Shwom, R.: Environmental values. Annual Review of Environment and Resources 30(1), 335–372 (2005)

[68] d'Inverno, M., Luck, M., Noriega, P., Rodriguez-Aguilar, J.A., Sierra, C.: Communicating open systems. Artificial Intelligence 186, 38–94 (2012)

[69] Domene, E., Saurí, D.: Urbanisation and Water Consumption: Influencing Factors in the Metropolitan Region of Barcelona. Urban Studies 43(9), 1605–1623 (2006)

[70] Dunn, G., Brown, R.R., Bos, J.J., Bakker, K.: Standing on the shoulders of giants: Understanding changes in urban water practice through the lens of complexity science. Urban Water Journal 14(7), 758–767 (2017)

[71] Edmonds, B.: Assessing the safety of (numerical) representation in social simulation. In: The 3rd European Social Simulation Association conference (ESSA 2005) (2005), `http://cfpm.org/cpmrep153.html`

[72] Edmonds, B.: Different modelling purposes. In: Edmonds, B., Meyer, R. (eds.) Simulating Social Complexity: A Handbook, pp. 38–58. Springer International Publishing, Cham (2017)

[73] Edmonds, B., Le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola Sales, C., Ormerod, P., Root, H., Squazzoni, F.: Different modelling purposes. Journal of Artificial Societies and Social Simulation 22(3), 6 (2019), `http://jasss.soc.surrey.ac.uk/22/3/6.html`

[74] Egidi, M.: Schumpeter's picture of economic and political institutions in the light of a cognitive approach to human behavior. Journal of Evolutionary Economics 27(1), 139–159 (2017)

[75] Epley, N., Tannenbaum, D.: Treating ethics as a design problem. Behavioral Science & Policy 3(2), 72–84 (2017)

[76] European Comission: Better Regulation Toolbox, [Online. Retrieved 2019, March 20] `https://ec.europa.eu/info/better-regulation-toolbox_en`

[77] European Commission: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions on the Roadmap to a Resource Efficient Europe (COM(2011) 0571 final) (2011)

[78] European Environment Agency: Water use at home (Infographic), [Online. Retrieved 2018, September 14] `https://www.eea.europa.eu/signals/signals-2018-content-list/infographic/water-use-at-home/view`

[79] EUROSTAT: Water statistics (2017), [Online. Retrieved 2018, December 11] `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Water_statistics`

[80] FAO: Water withdrawals by sectors, 2010 (2016), [Online. Retrieved 2017, September 6] `http://www.fao.org/nr/water/aquastat/tables/WorldData-Withdrawal_eng.pdf`

[81] FAO: The future of food and agriculture — Trends and challenges. Tech. rep., United Nations, Rome (2017)

[82] Federación Española de Municipios y Provincias (FEMP) & Asociación Española de Abastecimientos de Agua y Saneamiento (AEAS): Guía para la aplicación de Mecanismos de Acción Social en el sector del agua urbana *(in Spanish)*. Tech. rep., Madrid (2017)

[83] Feldman, S.: Structure and Consistency in Public Opinion: the Role of Core Beliefs and Values. American Journal of Political Science 32(2), 416–440 (1988)

[84] Fioramonti, L., Coscieme, L., Mortensen, L.F.: From gross domestic product to wellbeing: How alternative indicators can help connect the new economy with the Sustainable Development Goals. The Anthropocene Review (2019)

[85] Floridi, L. (ed.): The Onlife Manifesto, pp. 7–13. Springer International Publishing (2015)

[86] Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., Vayena, E.: AI4People — An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. Minds and Machines 28(4), 689–707 (2018)

[87] Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., Santos, J.I., del Olmo, R., López-Paredes, A., Edmonds, B.: Errors and artefacts in agent-based modelling. Journal of Artificial Societies and Social Simulation 12(1), 1 (2009)

[88] Galán, J.M., López-Paredes, A., Olmo, R.d.: An agent-based model for domestic water management in Valladolid metropolitan area. Water Resources Research 45(5) (2009)

[89] Galán Ordax, J.M.: Evaluación integradora de políticas de agua: modelado y simulación con sociedades artificiales de agentes. Ph.D. thesis, Universidad de Burgos (2007)

[90] Gilbert, G.N., Conte, R.: Artificial Societies: The Computer Simulation of Social Life. UCL Press, London (1995)

[91] Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., Narasimhan, K.P., Wilkinson, H.: Computational modelling of public policy: Reflections on practice. Journal of Artificial Societies and Social Simulation 21(1), 14 (2018)

[92] Gilbert, N., Troitzsch, K.: Simulation for the social scientist. McGraw-Hill Education (2005)

[93] Gleick, P., Iceland, C.: Water, Security and Conflict (August 2018), World Resources Institute

[94] Goren, P.: Core Principles and Policy Reasoning in Mass Publics: A Test of Two Theories. British Journal of Political Science 31(1), 159–177 (2001)

[95] Gorrieri, R.: Labeled Transition Systems, pp. 15–34. Springer (2017)

[96] Gotts, N.M., van Voorn, G.A., Polhill, J.G., de Jong, E., Edmonds, B., Hofstede, G.J., Meyer, R.: Agent-based modelling of socio-ecological systems: Models, projects and ontologies. Ecological Complexity (2018)

[97] Green, B.: The Smart Enough City: Putting Technology in Its Place to Reclaim Our Urban Future. Strong Ideas, MIT Press (2019)

[98] Greshko, M., Parker, L., Howard, B.C., Stone, D., Borunda, A., Gibbens, S.: A running list of how President Trump is changing environmental policy (2017, March 31), [Online. Retrieved 2019, March 15] `https://news.nationalgeographic.com/2017/03/how-trump-is-changing-science-environment/`

[99] Groenfeldt, D., Schmidt, J.J.: Ethics and water governance. Ecology and Society 18(1), 14 (2013)

[100] Grotenbreg, S., Altamirano, M.: Government facilitation of external initiatives: how dutch water authorities cope with value dilemmas. International Journal of Water Resources Development pp. 1–26 (2017)

[101] Guest, H.S.: Maslow's Hierarchy of Needs — the sixth level. The psychologist 27(12), 982–983 (2014)

[102] Guillaume Deffuant, Frédéric Amblard, G.W., Faure, T.: How can extremism prevail? A study based on the relative agreement interaction model. Journal of Artificial Societies and Social Simulation 5(4) (2002), `http://jasss.soc.surrey.ac.uk/5/4/1.html`

[103] Hamill, L., Gilbert, N.: Social circles: A simple structure for agent-based social network models. Journal of Artificial Societies and Social Simulation 12(2), 3 (2009), `http://jasss.soc.surrey.ac.uk/12/2/3.html`

[104] Hanel, P.H.P., Litzellachner, L.F., Maio, G.R.: An Empirical Comparison of Human Value Models. Frontiers in Psychology 9, 1643 (2018)

[105] Hansis, R.: Social acceptability in anthropology and geography. In: Defining Social Acceptability in Ecosystem Management, pp. 37–47. United States Department of Agriculture (1996)

[106] Heidari, S., Dignum, F., Jensen, M.: Simulation with values. In: Proceedings of the Social Simulation Conference 2018 (SSC2018) (In Press)

[107] Hibbert, C.: Simulation: the Bedrock of AI (2019), [Online. Retrieved 2019, January 20] `https://simudyne.com/resources/simulation-the-bedrock-of-ai/`

[108] Hitlin, S.: Values as the core of personal identity: Drawing links between two theories of self. Social psychology quarterly 66(2), 118–137 (2003)

[109] Hitlin, S.: Values, personal identity, and the moral self. In: Handbook of identity theory and research, pp. 515–529. Springer New York, New York (2011)

[110] Hitlin, S., Piliavin, J.A.: Values: Reviving a Dormant Concept. Annual Review of Sociology 30(1), 359–393 (2004)

[111] Hitlin, S., Pinkston, K.: Values, attitudes, and ideologies: explicit and implicit constructs shaping perception and action. In: DeLamater, J., Ward, A. (eds.) Handbook of Social Psychology, pp. 319–339. Springer Netherlands, Dordrecht (2013)

[112] HM Treasury (UK Government): The Magenta Book: Guidance for Evaluation (2011)

[113] Holtz, G., Pahl-Wostl, C.: An agent-based model of groundwater over-exploitation in the Upper Guadiana, Spain. Regional Environmental Change 12(1), 95–121 (2012)

[114] Hood, C.: A public management for all seasons? Public Administration 69(1), 3–19 (1991)

[115] Hoppe, R.: Heuristics for practitioners of policy design: Rules-of-thumb for structuring unstructured problems. Public Policy and Administration 33(4), 384–408 (2018)

[116] Hsieh, N.h.: Incommensurable values. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, spring 2016 edn. (2016)

[117] Hu, Y., Quinn, C.J., Cai, X., Garfinkle, N.W.: Combining human and machine intelligence to derive agents' behavioral rules for groundwater irrigation. Advances in Water Resources 109, 29–40 (2017)

[118] Huber, R., Bakker, M., Balmann, A., Berger, T., Bithell, M., Brown, C., Grêt-Regamey, A., Xiong, H., Le, Q.B., Mack, G., Meyfroidt, P., Millington, J., Müller, B., Polhill, J.G., Sun, Z., Seidl, R., Troost, C., Finger, R.: Representation of decision-making in European agricultural agent-based models. Agricultural Systems 167, 143–160 (2018)

[119] Inglehart, R.: The Silent Revolution in Europe: Intergenerational Change in Post-Industrial Societies. The American Political Science Review 65(4), 991–1017 (1971)

[120] Institut d'Estadística de Catalunya: Condicions de vida i hàbits de la població 2011 *(in Catalan)*, [Online. Retrieved 2018, October 15] `http://www.idescat.cat/pub/?id=ecvhp&n=6554&lang=en`

[121] Institute of Electrical and Electronics Engineers: Ethically aligned design (2017), [Online. Retrieved 2019, June 13] `https://ethicsinaction.ieee.org/`

[122] Instituto Nacional de Estadística: Censo de Población y Viviendas 2011 *(in Spanish)*, [Online. Retrieved 2018, October 15] `http://www.ine.es/censos2011_datos/cen11_datos_resultados.htm`

[123] Instituto Nacional de Estadística: Estadística sobre el Suministro y Saneamiento del Agua *(in Spanish)*, [Online. Retrieved 2019, September 18] `https://www.ine.es/prensa/essa_2016.pdf`

[124] Janssen, M., Wimmer, M.A.: Introduction to policy-making in the digital age. In: Janssen, M., Wimmer, M.A., Deljoo, A. (eds.) Policy Practice and Digital Science: Integrating Complex Systems, Social Simulation and Public Administration in Policy Research, pp. 1–14. Springer International Publishing, Cham (2015)

[125] Jasanoff, S., Wynne, B.: Science and decision making. In: Rayner, S., Malone, E.L. (eds.) Human choice and climate change, pp. 1–87. Battelle Press, Columbus, OH (1998)

[126] Jones, N.A., Ross, H., Lynam, T., Perez, P., Leitch, A.: Mental models: An interdisciplinary synthesis of theory and methods. Ecology and Society 16(1), 46 (2011)

[127] Jorgensen, B., Graymore, M., O'Toole, K.: Household water use behavior: An integrated model. Journal of Environmental Management 91(1), 227–236 (2009)

[128] Jørgensen, T.B., Bozeman, B.: Public values: An inventory. Administration & Society 39(3), 354–381 (2007)

[129] Jørgensen, T.B., Rutgers, M.R.: Public values: Core or confusion? introduction to the centrality and puzzlement of public values research. The American Review of Public Administration 45(1), 3–12 (2015)

[130] Kahneman, D., Tversky, A.: Choices, values, and frames. American Psychologist 39(4), 341 (1984)

[131] Kamateri, E., Panopoulou, E., Tambouris, E., Tarabanis, K., Ojo, A., Lee, D., Price, D.: A comparative analysis of tools and technologies for policy making. In: Janssen, M., Wimmer, M.A., Deljoo, A. (eds.) Policy Practice and Digital Science: Integrating Complex Systems, Social Simulation and Public Administration in Policy Research, pp. 125–156. Springer International Publishing, Cham (2015)

[132] Kaptein, M.: The Moral Entrepreneur: A New Component of Ethical Leadership. Journal of Business Ethics 156(4), 1135–1150 (2019)

[133] Keeney, R.L.: Value-Focused Thinking at British Columbia Hydro, pp. 342–371. Harvard University Press (1992)

[134] Keeney, R.: Value-Focused Thinking. Harvard University Press (1992)

[135] Kenter, J.O., O'Brien, L., Hockley, N., Ravenscroft, N., Fazey, I., Irvine, K.N., Reed, M.S., Christie, M., Brady, E., Bryce, R., Church, A., Cooper, N., Davies, A., Evely, A., Everard, M., Fisher, J.A., Jobstvogt, N., Molloy, C., Orchard-Webb, J., Ranger, S., Ryan, M., Watson, V., Williams, S.: What are shared and social values of ecosystems? Ecological Economics 111, 86–99 (2015)

[136] Kernaghan, K.: Integrating Values into Public Service: The Values Statement as Centerpiece. Public Administration Review 63(6), 711–719 (2003)

[137] Kiesling, E., Günther, M., Stummer, C., Wakolbinger, L.M.: Agent-based simulation of innovation diffusion: A review. Central European Journal of Operations Research 20(2), 183–230 (2012)

[138] Kitchin, R.: Thinking critically about and researching algorithms. Information, Communication & Society 20(1), 14–29 (2017)

[139] Knox, J.W., Kay, M.G., Weatherhead, E.K.: Water regulation, crop production, and agricultural water management — Understanding farmer perspectives on irrigation efficiency. Agricultural Water Management 108, 3–8 (2012)

[140] Konar, M., Garcia, M., Sanderson, M.R., Yu, D.J., Sivapalan, M.: Expanding the Scope and Foundation of Sociohydrology as the Science of Coupled Human-Water Systems. Water Resources Research 55(2), 874–887 (2019)

[141] Koutiva, I.: Agent-based modelling of domestic water demand for adaptive urban water management. Ph.D. thesis, National Technical University of Athens (2015)

[142] Lakoff, G.: Don't Think of an Elephant!: Know Your Values and Frame the Debate. Chelsea Green Publishing Company (2004)

[143] Lakoff, G.: Why it matters how we frame the environment. Environmental Communication 4(1), 70–81 (2010)

[144] Lakoff, G.: Why are so many people about to vote against their interests? (2010, 16 September), [Online. Retrieved 2019, February 06] `https://blogs.berkeley.edu/2010/09/16/why-are-so-many-people-about-to-vote-against-their-interests`

[145] Little, J.D.C.: Models and Managers: The Concept of a Decision Calculus. Management Science 16(8), B466–B485 (1970)

[146] Liu, J., Dietz, T., Carpenter, S.R., Folke, C., Alberti, M., Redman, C.L., Schneider, S.H., Ostrom, E., Pell, A.N., Lubchenco, J., Taylor, W.W., Ouyang, Z., Deadman, P., Kratz, T., Provencher, W.: Coupled human and natural systems. AMBIO: A Journal of the Human Environment 36(8), 639–649 (2007)

[147] López-Sánchez, M., Serramia, M., Rodrıguez-Aguilar, J., Morales, J., Wooldridge, M., Ansotegui, C.: Exploiting moral values to choose the right norms. In: Proceedings of the 1st Conference on Artificial intelligence, Ethics and Society, AIES. pp. 2–3 (2018)

[148] López-Gunn, E., Zorrilla, P., Prieto, F., Llamas, M.R.: Lost in translation? Water efficiency in Spanish agriculture. Agricultural Water Management 108, 83–95 (2012)

[149] López-Gunn, E.: *Agua Para Todos*: A new regionalist hydraulic paradigm in spain. Water Alternatives 2(3) (2009)

[150] López-Gunn, E., Mayor, B., Dumont, A.: Implications of the modernization of irrigation systems. In: De Stefano, L., Llamas, M.R. (eds.) Water, Agriculture and the Environment in Spain, chap. 19, pp. 241–255. CRC Press (2012)

[151] MacCarthaigh, M.: Public service values. Tech. rep., Institute of Public Administration Dublin (2008)

[152] Mair, D., Smillie, L., La Placa, G., Schwendinger, F., Raykovska, M., Pasztor, Z., Van Bavel, R.: Understanding our political nature: how to put knowledge and reason at the heart of political decision-making. Tech. rep., Joint Research Centre (2019)

[153] MAPAMA: Surfaces and annual productions of crops (in spanish) (2017), [Online. Retrieved 2017, June 15] `https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/superficies-producciones-anuales-cultivos/`

[154] MAPAMA: Agroclimatic information system for irrigation (in spanish) (ND), [Online. Retrieved 2017, June 15] `http://eportal.mapama.gob.es/websiar/NecesidadesHidricas.aspx`

[155] de Marchi, S., Page, S.E.: Agent-based models. Annual Review of Political Science 17(1), 1–20 (2014)

[156] Marcus, G., Davis, E.: Rebooting AI: Building Artificial Intelligence we can trust. Knopf Doubleday Publishing Group (2019)

[157] Martin-Ortega, J., Ferrier, R.C., Gordon, I.J., Khan, S., et al.: Water ecosystem services: A global perspective. UNESCO Publishing (2015)

[158] Mashhadi Ali, A., Shafiee, M.E., Berglund, E.Z.: Agent-based modeling to simulate the dynamics of urban water supply: Climate, population growth, and water shortages. Sustainable Cities and Society 28, 420–434 (2017)

[159] Maslow, A.H.: A theory of human motivation. Psychological review 50(4), 370–396 (1943)

[160] May, P.J.: Policy design and implementation. In: Peters, B., Pierre, J. (eds.) The SAGE Handbook of Public Administration, pp. 279–291. SAGE Publications, 2nd edn. (2012)

[161] Mazzucato, M.: The value of everything: Making and taking in the global economy. Hachette UK (2018)

[162] Meisch, S.P.: I Want to Tell You a Story: How Narrative Water Ethics Contributes to Re-theorizing Water Politics. Water 11(4), 631 (2019)

[163] Mercuur, R., Dignum, V., Jonker, C.: The use of values for modeling social agents. In: Quan Bai, Fenghui Ren, M.Z.T.I. (ed.) Proceedings of the 3nd International Workshop on Smart Simulation and Modelling for Complex Systems (2017)

[164] Millennium Ecosystem Assessment: Ecosystems and human well-being: Wetlands and water, vol. 5. Island Press, Washington, DC (2005)

[165] Milly, P.C., Betancourt, J., Falkenmark, M., Hirsch, R.M., Kundzewicz, Z.W., Lettenmaier, D.P., Stouffer, R.J.: Stationarity is dead: Whither water management? Science 319(5863), 573–574 (2008)

[166] Minsker, B., Baldwin, L., Crittenden, J., Kabbes, K., Karamouz, M., Lansey, K., Malinowski, P., Nzewi, E., Pandit, A., Parker, J., Rivera, S., Surbeck, C., Wallace, W.A., Williams, J.: Progress and recommendations for advancing performance-based sustainable and resilient infrastructure design. Journal of Water Resources Planning and Management 141(12), A4015006 (2015)

[167] Moody, R., Gerrits, L.: Values in computational models revalued. In: Janssen, M., Wimmer, M.A., Deljoo, A. (eds.) Policy Practice and Digital Science: Integrating Complex Systems, Social Simulation and Public Administration in Policy Research, pp. 205–219. Springer International Publishing, Cham (2015)

[168] Moore, M.H.: Creating public value. Harvard University Press (1995)

[169] Morais, C., de Moura, G.R.: The Psychology of Ethical Leadership in Organisations: Implications of Group Processes. Springer (2018)

[170] Moss, S., Edmonds, B.: Sociology and Simulation: Statistical and Qualitative Cross-Validation. American Journal of Sociology 110(4), 1095–1131 (2005)

[171] Munda, G.: On the use of Cost-Benefit Analysis and Multi-Criteria Evaluation in ex-ante Impact Assessment. Tech. Rep. EUR 28768 EN, Joint Research Center, Luxembourg (2017)

[172] Naranjo Chicharro, J.E.: Problems of irrigation modernisation (in Spanish) (2010), Communication in XII National Congress of Spain Irrigation Communities

[173] Nersessian, N.J., MacLeod, M.: Models and Simulations. In: Magnani, L., Bertolotti, T. (eds.) Springer Handbook of Model-Based Science. Springer (2017)

[174] Noriega, P., Padget, J., Verhagen, H., d'Inverno, M.: Towards a framework for socio-cognitive technical systems. In: Ghose, A., Oren, N., Telang, P., Thangarajah, J. (eds.) Coordination, Organizations, Institutions, and Norms in Agent Systems X. pp. 164–181. Lecture Notes in Computer Science 9372, Springer (2015)

[175] Noriega, P., Sabater-Mir, J., Verhagen, H., Padget, J., d'Inverno, M.: Identifying affordances for modelling second-order emergent phenomena with the wit framework. In: Sukthankar, G., Rodriguez, J.A. (eds.) AAMAS 2017 Workshops Visionary Papers. Springer International Publishing (2017)

[176] Noriega, P., Verhagen, H., d'Inverno, M., Padget, J.: A manifesto for conscientious design of hybrid online social systems. In: Coordination, Organizations, Institutions, and Norms in Agent Systems XII, pp. 60–78. Springer (2016)

[177] O'Brien, K.L., Wolf, J.: A values-based approach to vulnerability and adaptation to climate change. Wiley Interdisciplinary Reviews: Climate Change 1(2), 232–242 (2010)

[178] Oishi, S., Diener, E., Lucas, R.E.: The optimum level of well-being: Can people be too happy? In: Diener, E. (ed.) The Science of Well-Being, pp. 175–200. Springer Netherlands (2009)

[179] Ortega, M.V.: Collective management of irrigation in eastern Spain. Integration of new technologies and water resources. Ph.D. thesis, Universitat Politècnica de València (UPV), Spain (2015)

[180] Owens, S.: ‘Engaging the Public’: Information and Deliberation in Environmental Policy. Environment and Planning A: Economy and Space 32(7), 1141–1148 (2000)

[181] Pahl-Wostl, C.: Water Governance in the Face of Global Change. Springer International Publishing, Cham (2015)

[182] Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., Deadman, P.: Multi-agent systems for the simulation of land-use and land-cover change: A review. Annals of the Association of American Geographers 93(2), 314–337 (2003)

[183] Parks, L., Guay, R.P.: Personality, values, and motivation. Personality and Individual Differences 47(7), 675–684 (2009)

[184] Partington, R.: Inequality: is it rising, and can we reverse it? (2019, September 09), [Online. Retrieved 2019, November 12] `https://www.theguardian.com/news/2019/sep/09/inequality-is-it-rising-and-can-we-reverse-it`

[185] Pauwels, P.: Models in Architectural Design. In: Magnani, L., Bertolotti, T. (eds.) Springer Handbook of Model-Based Science. Springer (2017)

[186] Perello-Moragues, A., Noriega, P.: A Playground for the Value Alignment Problem. In: Martínez-Villaseñor, L., Batyrshin, I., Marín-Hernandez, A., Ponce, H. (eds.) Advances in Soft Computing. MICAI 2019. Lecture Notes in Computer Science, vol 11835. pp. 414–429. Springer International Publishing, Cham (2019)

[187] Perello-Moragues, A., Noriega, P.: Using Agent-Based Simulation to understand the role of values in policy-making. In: Verhagen, H., Borit, M., Bravo, G., Wijermans, N. (eds.) Advances in social simulation — Looking in the mirror. Springer (In Press)

[188] Perello-Moragues, A., Noriega, P., Padget, J., Verhagen, H.: Value-driven policy-making as a socio-cognitive technical system. In: First international workshop on socio-cognitive systems (SCS18) in the International Conference on Autonomous Agents and Multiagent Systems (AAMAS18) (2018)

[189] Perello-Moragues, A., Noriega, P., Poch, M.: Modelling contingent technology adoption in farming irrigation communities. Journal of Artificial Societies and Social Simulation 22(4), 1 (2019)

[190] Perello-Moragues, A., Noriega, P., Popartan, A., Poch, M.: Modelling policy shift advocacy. In: Proceedings of the Multi-Agent-Based Simulation Workshop (MABS) in the International Conference on Autonomous Agents and Multiagent Systems (AAMAS19) (In Press)

[191] Perello-Moragues, A., Noriega, P., Popartan, A., Poch, M.: On three ethical aspects involved in using agent-based social simulation for policy-making. In: Proceedings of the Social Simulation Conference 2019 (In Press)

[192] Perry, C.: ABCDE+F: A framework for thinking about water resources management. Water International 38(1), 95–107 (2013)

[193] Perry, C., Steduto, P., Allen, R.G., Burt, C.M.: Increasing productivity in irrigated agriculture: Agronomic constraints and hydrological realities. Agricultural Water Management 96(11), 1517–1524 (2009)

[194] Pfeiffer, L., Lin, C.Y.C.: Does efficient irrigation technology lead to reduced groundwater extraction? Empirical evidence. Journal of Environmental Economics and Management 67(2), 189–208 (2014)

[195] Van de Poel, I.: Values in engineering design. In: Meijers, A.W.M. (ed.) Handbook of the Philosophy of Science, pp. 973–1006. Elsevier (2009)

[196] Van de Poel, I.: Translating values into design requirements. In: Michelfelder, D.P., McCarthy, N., Goldberg, D.E. (eds.) Philosophy and Engineering: Reflections on Practice, Principles and Process, pp. 253–266. Springer Netherlands, Dordrecht (2013)

[197] van de Poel, I.: Conflicting values in design for values. In: van den Hoven, J., Vermaas, P.E., van de Poel, I. (eds.) Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains, pp. 89–116. Springer Netherlands, Dordrecht (2015)

[198] Pons, O., de la Fuente, A., Aguado, A.: The Use of MIVES as a Sustainability Assessment MCDM Method for Architecture and Civil Engineering Applications. Sustainability 8(5), 460 (2016)

[199] Prevos, P.: The invisible water utility: employee behaviour and customer experience in service-dominant logic. Ph.D. thesis, College of Arts, Social Sciences and Commerce, La Trobe University (2016)

[200] Primiero, G.: A Minimalist Epistemology for Agent-Based Simulations in the Artificial Sciences. Minds and Machines 29(1), 127–148 (2019)

[201] Region of Murcia: Regional agricultural statistics (2017), [Online. Retrieved 2017, June 15] `http://www.carm.es/web/pagina?IDCONTENIDO=1174&IDTIPO=100&RASTRO=c934$m`

[202] Regional Statistics Center of Murcia: Evolution of the surface and production according to type of crop (2017), [Online. Retrieved 2017, June 17] `http://econet.carm.es/inicio/-/crem/sicrem/PU_datosBasicos/sec49.html`

[203] Reiss, S.: Multifaceted nature of intrinsic motivation: The theory of 16 basic desires. Review of general psychology 8(3), 179–193 (2004)

[204] Rockström, J., Steffen, W., Noone, K., Åsa Persson, Chapin, F.S.I., Lambin, E., Lenton, T.M., Scheffer, M., Folke, C., Schellnhuber, H.J., Nykvist, B., de Wit, C.A., Hughes, T., van der Leeuw, S., Rodhe, H., Sörlin, S., Snyder, P.K., Costanza, R., Svedin, U., Falkenmark, M., Karlberg, L., Corell, R.W., Fabry, V.J., Hansen, J., Walker, B., Liverman, D., Richardson, K., Crutzen, P., Foley, J.: Planetary boundaries: Exploring the safe operating space for humanity. Ecology and Society 14(2) (2009)

[205] Rogers, E.M.: Diffusion of innovations. Free Press, New York, 4th edn. (1995)

[206] Rohan, M.J.: A Rose by Any Name? The Values Construct. Personality and Social Psychology Review 4(3), 255–277 (2000)

[207] Rokeach, M.: The nature of human values. Free Press (1973)

[208] Román-Cervantes, C.: Mercado de Tierras. In: Use and exploitation of the land in the region of Campo de Cartagena, (19th and 20th Centuries) *(In Spanish)*, pp. 197–304. Ministerio de Agricultura, Pesca y Alimentación, Madrid (1996)

[209] Roobavannan, M., Emmerik, T.H.M.v., Elshafei, Y., Kandasamy, J., Sanderson, M.R., Vigneswaran, S., Pande, S., Sivapalan, M.: Norms and values in sociohydrological models. Hydrology and Earth System Sciences 22, 1337–1349 (2018)

[210] Rosaria Conte, Giulia Andrighetto, M.C., Paolucci, M.: Emergent and immergent effects in complex social systems. In: Proceedings of AAAI Symposium, Social and Organizational Aspects of Intelligence. pp. 8–11 (2007)

[211] Russell, S.: Provably beneficial artificial intelligence. The Next Step: Exponential Life, BBVA-Open Mind (2017)

[212] Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall Press, 3rd edn. (2009)

[213] Saaty, T.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)

[214] Sabatier, P.A.: Theories of the policy process. Westview Press (1999)

[215] Sanchis-Ibor, C., Boelens, R., García-Mollá, M.: Collective irrigation reloaded. Re-collection and re-moralization of water management after privatization in Spain. Geoforum 87, 38–47 (2017)

[216] Saurí, D.: The decline of water consumption in Spanish cities: structural and contingent factors. International Journal of Water Resources Development 0(0), 1–17 (2019)

[217] Schreinemachers, P., Berger, T.: An agent-based simulation model of human–environment interactions in agricultural systems. Environmental Modelling & Software 26(7), 845–859 (2011)

[218] Schulz, C.: Governance-related values as dimensions of good water governance. Wiley Interdisciplinary Reviews: Water 0(0), e1322 (2018)

[219] Schwartz, S.H.: An Overview of the Schwartz Theory of Basic Values. Online Readings in Psychology and Culture 2(1) (2012)

[220] Schwartz, S., Cieciuch, J., Vecchione, M., Davidov, E., Fischer, R., Beierlein, C., Ramos, A., Verkasalo, M., Lönnqvist, J.E., Demirutku, K., Dirilen-Gumus, O., Konty, M.: Refining the theory of basic individual values. Journal of Personality and Social Psychology 103(4), 663–688 (2012)

[221] Schwartz, S.H.: Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In: Zanna, M.P. (ed.) Advances in Experimental Social Psychology, vol. 25, pp. 1–65. Academic Press (1992)

[222] Schwartz, S.H.: Are there universal aspects in the structure and contents of human values? Journal of Social Issues 50(4), 19–45 (1994)

[223] Schwartz, S.H.: Basic human values: Theory, methods, and application. Risorsa Uomo (2007)

[224] Schwartz, S.H.: Universalism Values and the Inclusiveness of Our Moral Universe. Journal of Cross-Cultural Psychology 38(6), 711–728 (2007)

[225] Schwartz, S.H., Bilsky, W.: Toward a universal psychological structure of human values. Journal of Personality and Social Psychology 53(3), 550–562 (1987)

[226] Schwartz, S.H., Bilsky, W.: Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. Journal of personality and social psychology 58(5), 878 (1990)

[227] Schwartz, S.H., Caprara, G.V., Vecchione, M.: Basic Personal Values, Core Political Values, and Voting: A Longitudinal Analysis. Political Psychology 31(3), 421–452 (2010)

[228] Schwarz, N., Ernst, A.: Agent-based modeling of the diffusion of environmental innovations — An empirical approach. Technological Forecasting and Social Change 76(4), 497–511 (2009)

[229] SCRATS: Rates for Tajo-Segura transfers (in Spanish) (2017), [Online. Retrieved 2017, August 21] `http://www.scrats.es/tarifas-vigentes.html`

[230] Sekera, J.: The public economy: Understanding government as a producer. Tech. Rep. IIPP WP 2020-01, UCL Institute for Innovation and Public Purpose, London (2020)

[231] Sengupta, R., Lant, C., Kraft, S., Beaulieu, J., Peterson, W., Loftus, T.: Modeling Enrollment in the Conservation Reserve Program by Using Agents within Spatial Decision Support Systems: An Example from Southern Illinois. Environment and Planning B: Planning and Design 32(6), 821–834 (2005)

[232] Sharp, L.: Reconnecting people and water: public engagement and sustainable urban water management. Routledge (2017)

[233] Shindler, B.A., Brunson, M., Stankey, G.H.: Social acceptability of forest conditions and management practices: a problem analysis. Tech. Rep. PNW-GTR-537, United States Department of Agriculture, Portland (USA) (2002)

[234] Sierra, C., Osman, N., Noriega, P., Sabater-Mir, J., Perello-Moragues, A.: Value alignment: A formal approach. In: Responsible Artificial Intelligence Agents Workshop (RAIA) in the International Conference on Autonomous Agents and Multiagent Systems (AAMAS) (2019)

[235] Simon, H.A.: Bounded rationality. In: Eatwell, J., Milgate, M., Newman, P. (eds.) Utility and Probability, pp. 15–18. Palgrave Macmillan UK (1990)

[236] Simon, H.A.: The sciences of the artificial. MIT Press, 3rd edn. (1996)

[237] Simon, H.A.: Fact and Value in Decision-making. In: Administrative Behavior: A study of decision-making processes in administrative organization. The Free Press, 4th edn. (1997)

[238] Sirdeshmukh, D., Singh, J., Sabol, B.: Consumer Trust, Value, and Loyalty in Relational Exchanges. Journal of Marketing 66(1), 15–37 (2002)

[239] Sivapalan, M., Konar, M., Srinivasan, V., Chhatre, A., Wutich, A., Scott, C.A., Wescoat, J.L., Rodríguez-Iturbe, I.: Socio-hydrology: Use-inspired water sustainability science for the Anthropocene. Earth's Future 2(4), 225–230 (2014)

[240] Smil, V.: Nitrogen in crop production: An account of global flows. Global Biogeochemical Cycles 13(2), 647–662 (1999)

[241] Spring, A.: Sydney facing first water restrictions in a decade as drought grips New South Wales (2019, May 28), [Online. Retrieved 2019, May 31] `https://www.theguardian.com/australia-news/2019/may/28/sydney-facing-first-water-restrictions-in-a-decade-as-drought-grips-new-south-wales`

[242] Sprong, S., Jetten, J., Wang, Z., Peters, K., Mols, F., Verkuyten, M., Bastian, B., Ariyanto, A., Autin, F., Ayub, N., Badea, C., Besta, T., Butera, F., Costa-Lopes, R., Cui, L., Fantini, C., Finchilescu, G., Gaertner, L., Gollwitzer, M., Ángel Gómez, González, R., Hong, Y.Y., Jensen, D.H., Jasinskaja-Lahti, I., Karasawa, M., Kessler, T., Klein, O., Lima, M., Mégevand, L., Morton, T., Paladino, P.,

Polya, T., Renvik, T.A., Ruza, A., Shahrazad, W., Shama, S., Smith, H.J., Torres, A.R., van der Bles, A.M., Wohl, M.J.A.: "Our Country Needs a Strong Leader Right Now": Economic Inequality Enhances the Wish for a Strong Leader. Psychological Science 0(0) (2019)

[243] Srinivasan, V., Sanderson, M., Garcia, M., Konar, M., Blöschl, G., Sivapalan, M.: Prediction in a socio-hydrological world. Hydrological Sciences Journal 62(3), 338–345 (2017)

[244] Starr, P.: Seductions of sim: Policy as a simulation game. American Prospect 17, 19–29 (1994)

[245] Steduto, P., Hsiao, T.C., Fereres, E., Raes, D.: Crop yield response to water. United Nations (2012), FAO Paper 66

[246] Stephen Cranefield, Michael Winikoff, V.D.F.D.: No pizza for you: Value-based plan selection in bdi agents. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17. pp. 178–184 (2017)

[247] Stewart, J.: Public Policy Values. Palgrave Macmillan UK, London (2009)

[248] Sun, Z., Lorscheid, I., Millington, J.D., Lauf, S., Magliocca, N.R., Groeneveld, J., Balbi, S., Nolzen, H., Müller, B., Schulze, J., Buchmann, C.M.: Simple or complicated agent-based models? A complicated issue. Environmental Modelling & Software 86, 56–67 (2016)

[249] Susskind, J.: Future Politics: Living Together in a World Transformed by Tech. Oxford University Press (2018)

[250] Susskind, R.: AI, work and 'outcome-thinking' (2018, 1 November), [Online. Retrieved 2019, April 24] `https://www.thebritishacademy.ac.uk/publications/british-academy-review/34/ai-work-and-outcome-thinking`

[251] Symons, J., Alvarado, R.: Epistemic Entitlements and the Practice of Computer Simulation. Minds and Machines 29(1), 37–60 (2019)

[252] Teze, J.C., Perello-Moragues, A., Godo, L., Noriega, P.: Practical reasoning using values: an argumentative approach based on a hierarchy of values. Annals of Mathematics and Artificial Intelligence 87(3), 293–319 (2019)

[253] Thacher, D., Rein, M.: Managing value conflict in public policy. Governance 17(4), 457–486 (2004)

[254] Tortajada, C., Biswas, A.K.: The rapidly changing global water management landscape. International Journal of Water Resources Development 33(6), 849–851 (2017)

[255] Treviño, L.K., Weaver, G.R., Reynolds, S.J.: Behavioral ethics in organizations: A review. Journal of Management 32(6), 951–990 (2006)

[256] UK Cabinet Office: White Paper: Modernising Government (1999)

[257] UNDESA: International Decade for Action WATER FOR LIFE 2005–2015: Human right to water (2014, May 29), [Online. Retrieved 2019, November 05] `https://www.un.org/waterforlifedecade/human_right_to_water.shtml`

[258] UNESCO: Water, a shared responsibility. Tech. rep., United Nations, Paris (2006)

[259] United Nations: World Urbanization Prospects: The 2018 Revision (2018), Working report N° ESA/P/WP.252

[260] United Nations Conference on Housing and Sustainable Urban Development: Habitat III Issue Papers (2017)

[261] Van Dyke Parunak, H., Savit, R., Riolo, R.L.: Agent-Based Modeling vs. Equation-Based Modeling: A Case Study and Users' Guide. In: Sichman, J.S., Conte, R., Gilbert, N. (eds.) Multi-Agent Systems and Agent-Based Simulation. pp. 10–25. Springer Berlin Heidelberg (1998)

[262] Van Leeuwen, C.J., Frijns, J., van Wezel, A., van de Ven, F.H.M.: City blueprints: 24 indicators to assess the sustainability of the urban water cycle. Water Resources Management 26(8), 2177–2197 (2012)

[263] Vanhée, L., Ferber, J., Dignum, F.: Agent-Based Evolving Societies. In: Kamiński, B., Koloch, G. (eds.) Advances in Social Simulation. pp. 283–294. Springer Berlin Heidelberg (2014)

[264] Veale, M.: A critical take on the policy recommendations of the eu high-level expert group on artificial intelligence. LawArXiv (2019)

[265] Wagner, B.: Ethics as an escape from regulation. From "ethics-washing" to ethics-shopping?, pp. 84–89. Amsterdam University Press (2018)

[266] Watts, J.: Cape Town faces Day Zero: what happens when the city turns off the taps? (2018, February 03), [Online. Retrieved 2019, March 20] `https://www.theguardian.com/cities/2018/feb/03/day-zero-cape-town-turns-off-taps`

[267] Van der Weide, T.L., Dignum, F., Meyer, J.C., Prakken, H., Vreeswijk, G.A.W.: Practical reasoning using values: Giving meaning to values. In: Proceedings of the 6th International Conference on Argumentation in Multi-Agent Systems. pp. 79–93. ArgMAS'09, Springer-Verlag, Berlin, Heidelberg (2010)

[268] Wiek, A., Larson, K.L.: Water, People, and Sustainability—A Systems Framework for Analyzing and Assessing Water Governance Regimes. Water Resources Management 26(11), 3153–3171 (2012)

[269] de Wildt, T., Chappin, E., van de Kaa, G., Herder, P., van de Poel, I.: Conflicting values in the smart electricity grid a comprehensive overview. Renewable and Sustainable Energy Reviews 111, 184–196 (2019)

[270] Witesman, E., Walters, L.: Public Service Values: A new approach to the study of motivation in the public sphere. Public Administration 92(2), 375–405 (2014)

[271] Witesman, E.M., Walters, L.C.: Modeling public decision preferences using context-specific value hierarchies. The American Review of Public Administration 45(1), 86–105 (2015)

[272] Witesman, O.F.: Universality in Public Service Values: A Methodology for the Comparison of Value Structures and Content. Ph.D. thesis, School of Public and Environmental Affairs, Indiana University (2017)

[273] Wooldridge, M.: An Introduction to MultiAgent Systems. Wiley Publishing, 2nd edn. (2009)

[274] World Health Organisation: Nitrate and nitrite in drinking-water (2011), background document for development of WHO Guidelines for Drinking-water Quality

[275] Zhang, H., Vorobeychik, Y., Letchford, J., Lakkaraju, K.: Data-driven agent-based modeling, with application to rooftop solar adoption. Autonomous Agents and Multi-Agent Systems 30(6), 1023–1049 (2016)