This research was conducted within the research project “Sustainable Water Infrastructure Planning”, http://www.eawag.ch/forschung/sww/gruppen/swip/index_EN. It was funded within the National Research Programme 61 on “Sustainable Water Management” by the Swiss National Science Foundation (http://www.nrp61.ch).

Tackling uncertainty in multi-criteria decision analysis – An application to water supply infrastructure planning

Lisa Scholten [a*], Nele Schuwirth [a], Peter Reichert [a], Judit Lienert [a]

[a] *Eawag, Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, P.O. Box 611, CH-8600 Duebendorf, Switzerland*

* Corresponding author. E-Mail: lisa.scholten@eawag.ch

Published in European Journal of Operational Research (2014)

View at publisher: *http://www.sciencedirect.com/science/article/pii/S0377221714007838*

Tackling uncertainty in multi-criteria decision analysis – An application to water supply infrastructure planning

Lisa Scholten[a,b,*)], Nele Schuwirth[a)], Peter Reichert[a,c)], Judit Lienert[a)]

a) Eawag: Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, P.O.Box 611, CH-8600 Dübendorf, Switzerland

b) ETH Zürich, Institute of Environmental Engineering, CH-8093 Zürich, Switzerland

c) ETH Zürich, Institute of Biochemistry and Pollutant Dynamics (IBP), CH-8092 Zürich, Switzerland

* corresponding author:
Eawag: Swiss Federal Institute of Aquatic Science and Technology, Überlandstrasse 133, P.O. Box 611, CH-8600 Dübendorf, Switzerland.
Tel.: +41 58 765 5590; Fax: +41 58 765 5028.
lisa.scholten@eawag.ch

Abstract

We present a novel approach for practically tackling uncertainty in preference elicitation and predictive modeling to support complex multi-criteria decisions based on multi-attribute utility theory (MAUT). A simplified two-step elicitation procedure consisting of an online survey and face-to-face interviews is followed by an extensive uncertainty analysis. This covers uncertainty of the preference components (marginal value and utility functions, hierarchical aggregation functions, aggregation parameters) and the attribute predictions. Context uncertainties about future socio-economic developments are captured by combining MAUT with scenario planning. We perform a global sensitivity analysis (GSA) to assess the contribution of single uncertain preference parameters to the uncertainty of the ranking of alternatives. This is exemplified for sustainable water infrastructure planning in a case study in Switzerland. We compare eleven water supply alternatives ranging from conventional water supply systems to novel technologies and management schemes regarding 44 objectives. Their performance is assessed for four future scenarios and ten stakeholders from different backgrounds and decision-making levels. Despite uncertainty in the ranking of alternatives, potential best and worst solutions could be identified. We demonstrate that *a priori* assumptions such as linear value functions or additive aggregation can result in misleading recommendations, unless thoroughly checked during preference elicitation and modeling. We suggest GSA to focus elicitation on most sensitive preference parameters. Our GSA results indicate that output uncertainty can be considerably reduced by additional elicitation of few parameters, e.g. the overall risk attitude and aggregation functions at higher-level nodes. Here, rough value function elicitation was sufficient, thereby substantially reducing elicitation time.

Keywords

Decision analysis, Uncertainty modelling and global sensitivity analysis, multi-attribute utility theory, preference elicitation, water infrastructure planning

1 Introduction

1.1 Consideration of uncertainty in MAUT applications

Over the past decade, the number of applications of multi-criteria decision analysis (MCDA) and more specifically, multi-attribute utility theory (MAUT) and multi-attribute value theory (MAVT) (e.g. Keeney, 1982; Keeney & Raiffa, 1993), has considerably increased in the environmental sciences (Ananda & Herath, 2009; Huang, Keisler, & Linkov, 2011). This is also the case in other disciplines (Wallenius, et al., 2008). In MAUT applications, strong simplifying assumptions are often made to keep elicitation and modeling of preferences feasible given the available resources. Common simplifications are a) the choice of additive MAUT models (Hajkowicz, 2008; Hyde, Maier, & Colby, 2005; Joubert, Stewart, & Eberhard, 2003), b) use of linear marginal value functions (Raju & Vasan, 2007; Weber, 1987), c) assumption of risk neutrality, as well as d) neglecting uncertainty of model parameters (e.g. "weights"), attributes, and boundary conditions such as socio-economic change (Hyde, Maier, & Colby, 2004; Martin, Bender, & Shields, 2000; Torrance, et al., 1996). The reasons are manifold, e.g. higher model comprehensibility for decision makers, time constraints, and the need for cognitively tiring repetitive assessments (Karvetski, Lambert, & Linkov, 2009a; Stewart, 1995), but often remain undisclosed. Although the necessity of a systematic consideration of uncertainty has been widely acknowledged in theory (e.g. Butler, Jia, & Dyer, 1997; Durbach & Stewart, 2011, 2012b; French, 2003; Kangas & Kangas, 2004; Keeney & Raiffa, 1993; Stewart, 1995, 2005), it is commonly not considered in practice.

1.2 Sources of uncertainty

Different sources of uncertainty in MCDA are discussed in the literature. These cover uncertainties arising from (1) problem framing and structuring, (2) attribute prediction, and also (3) components of the preference model, i.e. in the case of MAVT and MAUT:(3a) the choice of hierarchical aggregation functions, (3b) the form of the marginal value / utility functions, and (3c) the corresponding aggregation parameters ("weights"). Furthermore, many of the commonly used preference elicitation techniques lack robustness towards biases (Bleichrodt, Pinto, & Wakker, 2001; Borcherding, Eppel, & von Winterfeldt, 1991; Morton & Fasolo, 2009; Weber & Borcherding, 1993), constituting an additional source of uncertainty.

By using the word "uncertainty" in this paper, we make no distinction between uncertainties elsewhere referred to as *risk* (known cause-effect, probabilistically quantifiable), *uncertainty* (known cause-effect, not probabilistically quantifiable), and *ignorance* ("deep uncertainty", unknown cause-effect, not quantifiable). Other classifications distinguish between *aleatory uncertainty* (due to randomness, see *risk*) and *epistemic uncertainty* (due to lack of knowledge, sometimes quantifiable). Instead, we use the term *uncertainty* when referring to "knowledge gaps or ambiguities that affect our ability to understand the consequences of decisions" (Gregory, et al., 2012, p.127), i.e. the way it is used in common language.

(1) Problem framing and structuring. Problem framing and structuring concerns the definition of the decision problem and boundary conditions, a stakeholder analysis to establish participation, and the development of the system of objectives and a set of alternatives for evaluation (Belton & Stewart, 2002; Keeney, 1982). Uncertainties arising from problem structuring are hardly quantifiable. People arrive at different decisions for the same problem dependent on the problem framing (Belton & Stewart, 2002; Morton & Fasolo, 2009). Different hierarchical structuring of the same system of objectives has been shown to affect the assigned weights (due to "splitting bias", e.g. Weber & Borcherding, 1993). Additionally, the number of identified fundamental objectives is linked to how well decision makers are supported during the formulation of fundamental objectives (Bond, Carlson, & Keeney, 2008, 2010). Thorough structuring is thus indispensable. An overview of structuring methods is given in e.g. Belton and Stewart (2010) and Franco and

Montibeller (2011). A growing trend in MCDA is to address uncertainties about future framework boundary conditions that are beyond the influence of decision makers with scenario analysis (e.g. Goodwin & Wright, 2001; Montibeller, Gummer, & Tumidei, 2006; Stewart, French, & Rios, 2013).

(2) Attribute prediction. The sources of uncertainty about the attribute levels of each decision alternative depend on the assessment process. Uncertainty can arise from the imprecision of quantitative elicitation and formulation of expert estimates which is prone to biases (Ayyub, 2001; Cooke, 1991; Kynn, 2008; O'Hagan, et al., 2006). It can also stem from the uncertainty of model predictions such as uncertainty of model input / structure / parameters (see e.g. French, 1995; Refsgaard, van der Sluijs, Højberg, & Vanrolleghem, 2007; Walker, et al., 2003).

(3) Hierarchical aggregation function. The multi-attribute value or utility function is typically structured hierarchically (see later example, Figure 1). The value or utility of the main objective depends on lower-level utility or value functions. These may directly depend on the attributes ("marginal utility or value functions") or indirectly through intermediate aggregation functions. The uncertainty about the hierarchical aggregation function is governed by the lack of knowledge about which independence conditions are satisfied by the decision maker's preferences (Eisenführ, Weber, & Langer, 2010; Keeney & Raiffa, 1993), and the precision of other aggregation model parameters. The additive, multiplicative, and multi-linear models are presented in Keeney and Raiffa (1993). The first requires *mutual preferential independence*, *additive independence*, and either *difference independence* (for values) or *mutual utility independence* (for utilities) to hold (Eisenführ, et al., 2010). The second model does not require additive independence. The third model requires the weakest assumptions, but easily becomes infeasible due to non-identifiability of its parameters (Stewart, 2005). Other less common models are the Cobb-Douglas model (i.e. the weighted geometric mean, originally suggested as a production function but later also used in the current context; Cobb & Douglas, 1928), minimum-models, or mixtures of these (e.g. Langhans, Lienert, Schuwirth, & Reichert, 2013; Langhans, Reichert, & Schuwirth, 2014; Schuwirth, Reichert, & Lienert, 2012).

(4) Marginal ("single-attribute") value or utility functions. Uncertainty about the shape of value and utility functions also arises from the imprecision of preferences, as well as inconsistencies and elicitation biases. Following von Neumann and Morgenstern (1947, in Eisenführ et al., 2010) and Dyer and Sarin (1979), we differentiate between (measurable) value functions and (ordinal) utility functions. Value functions describe preferences regarding sure attribute outcomes. Utility functions are used to rank "risky" attribute outcomes (the uncertainty of which is quantifiable by probability distributions). Utility functions are either directly elicited (Hershey & Schoemaker, 1985; Wakker & Deneffe, 1996) or obtained from converting value functions to utility functions given a specific intrinsic risk attitude (Dyer & Sarin, 1982). Again, several biases are known. For assigning values: *scope insensitivity* and *reference point effects* (e.g. Morton & Fasolo, 2009), and for the assessment of utilities (Bleichrodt, et al., 2001; Cox, Sadiraj, Vogt, & Dasgupta, 2012; Eisenführ, et al., 2010): *non-linear weighting of probabilities* (Kahneman & Tversky, 1979), *ambiguity aversion* (Ellsberg paradox; Ellsberg, 1961), and *certainty effects* (Allais paradox; Allais, 1953). In the absence of bias-free elicitation methods, some have questioned the use of expected utility theory (e.g. Abdellaoui, Bleichrodt, & Paraschiv, 2007; Cox, et al., 2012; Rabin, 2000; Schmidt, Starmer, & Sugden, 2008). Others developed approaches to correct for biases (Bleichrodt, et al., 2001) or simply accept some degree of descriptive deviation from theory in prescriptive decision analyses (e.g. French, 2003; Stewart, 2005).

(5) Aggregation parameters ("weights"). Uncertainty and imprecision of the weights are related to the articulated accuracy and consistency of judgments (Jessop, 2011). The elicitation of weights is prone to biases, such as the *splitting bias*, *range effect*, and *hierarchical effects* (Morton & Fasolo, 2009; Weber & Borcherding, 1993).

Comparing four weight elicitation methods, Borcherding et al. (1991) judge none to be internally more consistent or less biased than the others, and suggest doing more consistency checks. Mustajoki et al. (2005) and Jessop (2011) argue that the assumption of exact weights imposes a precision not represented by the stakeholder's preferences and recommend using imprecise or interval weights instead. Using imprecise weights also reduces inconsistencies within and between elicitation methods. Hierarchical elicitation (e.g. Pöyhönen, Vrolijk, & Hämäläinen, 2001) and *ex post* corrections (Jacobi & Hobbs, 2007) have been suggested to minimize the splitting bias.

1.3 Uncertainty and sensitivity analysis

Although often interchangeably used, the term *uncertainty analysis* refers to the quantification of model output uncertainty through propagation of uncertainty of model parameters and inputs (French, 2003), and *sensitivity analysis* to "the study of how uncertainty in the output […] can be apportioned to different sources of uncertainty in the model input" (Saltelli, Tarantola, Campolongo, & Ratto, 2004). Global sensitivity analysis (GSA) allows inputs to vary according to a given probability distribution, whereas local sensitivity analysis (LSA) uses a linearization of the model at a pre-defined point in parameter space (Saltelli, 2008). Uncertainty and sensitivity analyses address a range of modeling-related questions (e.g. French, 2003; Saltelli, 2008). Two of them are of particular interest to decision making: (1) How does the ranking of alternatives change, given the uncertainty of preference model inputs and (2) how strong is the influence of individual factors (to focus elicitation and modeling on reducing uncertainty that matters)?

In MAVT and MAUT, uncertainty and local sensitivity analyses are much more commonly performed than global sensitivity analyses (Gómez Delgado & Bosque Sendra, 2004; Saltelli, Ratto, Tarantola, & Campolongo, 2006; A. Saltelli, S. Tarantola, & K. Chan, 1999a). GSA has been suggested to support decision makers in the analysis of results from MCDA studies (Mustajoki, Hämälainen, & Lindstedt, 2006; Saltelli, et al., 1999a), but only applied in few cases (e.g. solid waste management: Gómez Delgado & Tarantola, 2006). The vast majority of available uncertainty and sensitivity analyses focusses on the uncertainty of the weights (e.g. Butler, et al., 1997; Hyde, et al., 2005; Jessop, 2011; Jiménez, Ríos-Insua, & Mateos, 2006; Mustajoki, 2012; Mustajoki, et al., 2006; Raju & Pillai, 1999) or a combination of aggregation parameters and attributes (e.g. Gómez Delgado & Bosque Sendra, 2004; Gómez Delgado & Tarantola, 2006; Hyde, et al., 2004; Saltelli, et al., 1999a). Zhou and Ang (2009) consider weights and two multi-attribute aggregation methods. Simulation studies by Stewart (2005) and Durbach and Stewart (2009, 2012a) assess the impact of hierarchical value und utility model simplifications under different marginal utility curvatures, degrees of imprecision in preference statements, and attributes among other aspects. Schuwirth et al. (2012) perform a LSA over changes of the weights, marginal value functions, risk attitudes for conversion to utilities, and the attributes. Another methodology for tackling uncertainty of the weights and marginal utility function curvatures is "Stochastic Multiobjective Acceptability Analysis" (SMAA; see e.g. Lahdelma, Hokkanen, & Salminen, 1998; Lahdelma & Salminen, 2012). SMAA is a simulation approach for determining which preference combinations would lead specific alternatives to rank best without requiring the decision makers' preferences to be known. The model structure only allows compensatory (additive) aggregation and risk neutral (value functions identical to utility functions) preferences.

1.4 Application of MAUT to water supply infrastructure planning

The planning of urban water supply infrastructures is an ideal application field for MAUT because it not only involves many, conflictive objectives and stakeholders, but also because of the high interactions with other systems, its long asset life times, and uncertain future development of main drivers of its performance. Urban water supply infrastructures in industrialized countries are mainly centralized treatment and piped distribution systems, which ensure a continuous supply of drinking water for households, industries, businesses, and public use (e.g. street-cleaning, public green space). They are facing a number of dynamic challenges such as urbanization and population development, aging and need of rehabilitation, climate variability, as well as a highly dynamic socio-economic and socio-political environment (Ferguson, Frantzeskaki, & Brown, 2013; Ruth, Bernier, Jollands, & Golubiewski, 2007; Sharma, Burn, Gardner, & Gregory, 2010). For a thorough planning of these water infrastructures, long-term changes and large uncertainties of drivers such as water availability, water demand, population and spatial development, and economic development need to be considered. Technically, transitions to more decentralized infrastructures (e.g. rainwater harvesting, or water treatment and reuse in households) are suggested to ensure flexible adaptation to future changes and increase sustainability (Sharma, et al., 2010; Wong & Brown, 2009). Additionally, alternative forms of utility governance can be chosen, e.g. regionalization or (partial) privatization to achieve higher efficiency and professionalism (Dominguez, Worch, Markard, Trujillo Alvarez, & Gujer, 2009; Lieberherr, Klinke, & Finger, 2012). In contrast to this, the reality of today's water infrastructure planning is often judged inflexible, narrow-minded, and negligent of future uncertainties, broader goals, important stakeholders, and alternative paths of action (Ashley, et al., 2008; Dominguez, et al., 2009; Economides, 2012; Ferguson, et al., 2013; Störmer, et al., 2009).

1.5 Aim of the study and main research questions

The objective of this paper is to show how to practically tackle uncertainty in elicitation and modeling of MAUT preferences. Our approach is developed and tested in a case study on sustainable water infrastructure planning in Switzerland. It is part of a larger study on water supply and sanitation planning introduced in Lienert et al. (2014a) . We use this case study to exemplify our approach, and present the results for water supply. This includes the elicitation of preferences of ten stakeholders, which were selected based on an earlier stakeholder analysis (Lienert, Schnetzer, & Ingold, 2013). To address the challenges of long-term infrastructure planning under uncertainty, the MCDA is combined with scenario planning. The study is guided by three main questions:

1. How can multiple sources of uncertainty in MAUT be comprehensively considered during elicitation and analysis of preferences?

2. Which uncertain preference parameters contribute most to the overall uncertainty of the ranking of alternatives, and how does this contribution change under different modeling assumptions?

3. What are the stakeholders' preferences regarding "good water supply infrastructure", and which water supply alternatives can be recommended given different future scenarios?

The case study and methods are presented in section 2. Of the above mentioned sources of uncertainty, the sources from (2) to (5) are quantitatively described. The uncertainty from framing and structuring (1), was considered by systematic structuring and framing within individual interviews and workshops (see Lienert, et al., 2013; Lienert, et al., 2014a), and preference elicitation including consistency checks. In section 3 we present the elicited preferences, attribute predictions, and resulting rankings of alternatives including uncertainty for four future scenarios. The results of the global sensitivity analysis for one exemplary stakeholder are shown in section 4. The results from 3 and 4 are discussed in section 5 and conclusions are drawn in section 6.

2 Material and methods

2.1 Case study "Mönchaltorfer Aa"

The "Mönchaltorfer Aa" region is a rural area near Zurich, Switzerland. Four municipalities (approx. 24'200 inhabitants) and five local water suppliers participated in the case study. Water infrastructures are either run by municipalities or cooperatives. Part of the water is imported from a regional cooperative. Despite an overall perception of high levels of service, supply security, and good water quality, some doubts about the long-term planning of the water supply system prevail.

Stakeholder identification

Lienert et al. (2013) identified 41 important actors for water and wastewater infrastructure planning, 29 of which are either shared between both sectors or are relevant to water supply only. Out of these, ten were selected to participate in the MCDA, eight of which were nominated based on their importance for water supply infrastructure planning (SH1–8), see supporting information (SI; section SI1). To ensure a better balance, and because of their importance for long-term legislative and political changes, we also included two stakeholders from the national level, although these were judged less important for local planning processes (SH9 and SH10). Detailed information on stakeholder classification is also given in Lienert et al. (2013). Together, they represent different entities and decision-making levels, summarized in Table 1:

Table 1: Participants of the MCDA. SH = stakeholder

No.	Name	Entity/responsibility	Level
SH1	Municipal underground engineer	Municipal representative in charge of underground engineering works	Local
SH2	Operating staff	Responsible for the technical functioning and monitoring of the water supply system	Local
SH3	Local water supply cooperative	Representative of water provider and operator of water infrastructures	Local
SH4	Municipal administration & finance	Municipal representative in charge of water supply services and finance	Local
SH5	Engineering consultant	Private consultant in charge of practical water infrastructure planning and technical dimensioning of water infrastructures	Regional
SH6	Regional water supply cooperative	Representative of regional water supply cooperative which delivers water from sources outside the case study region and operates transport infrastructures for altogether fourteen water utilities. The five case study utilities are shareholders of the regional cooperative.	Regional
SH7	Cantonal environmental protection agency	Representative of cantonal environmental protection authority which monitors and regulates the quality and use of water resources and approves of water infrastructure planning; implements national and cantonal water-related legislation	Cantonal
SH8	Cantonal (water) quality laboratory	Representative of cantonal authority which controls and approves water quality (among other products)	Cantonal
SH9	Swiss gas and water industry association	Representative of Swiss gas and water industry association which trains and accredits technical operating staff and designs and publishes relevant technical guidelines for water supply	National
SH10	National environmental protection agency	Representative of national authority which monitors the use and quality of water resources on the national scale, implements national environmental laws and regulation, and prepares political decisions	National

Objectives hierarchy and attributes

The objectives hierarchy in Figure 1 was developed in individual interviews and a stakeholder workshop (Lienert, et al., 2014a). The overall objective of achieving a "good water supply infrastructure" constitutes five fundamental objectives: "high intergenerational equity", "high resources and groundwater protection", "good water supply", "high social acceptance", and "low costs". These are divided into sub-objectives which are directly measured by a corresponding attribute, except "good water supply" , Figure 1. The sub-objectives of "good water supply" are further divided into "good drinking water (dw) supply", "good household water (hw) supply", and "good firefighting water (ffw) supply", since these are separately supplied in some alternatives. They are characterized by the same sub-objectives concerning water quantity, reliability, and quality. The latter is not considered for "good ffw supply", because water quality is irrelevant for firefighting. The attributes and their assessment are explained in Table SI2.1 (SI). Some attribute ranges had to be chosen generously to ensure that the predictions for the decision alternatives (incl. uncertainty) were covered, because the final predictions were still missing at the time of the MCDA interviews.

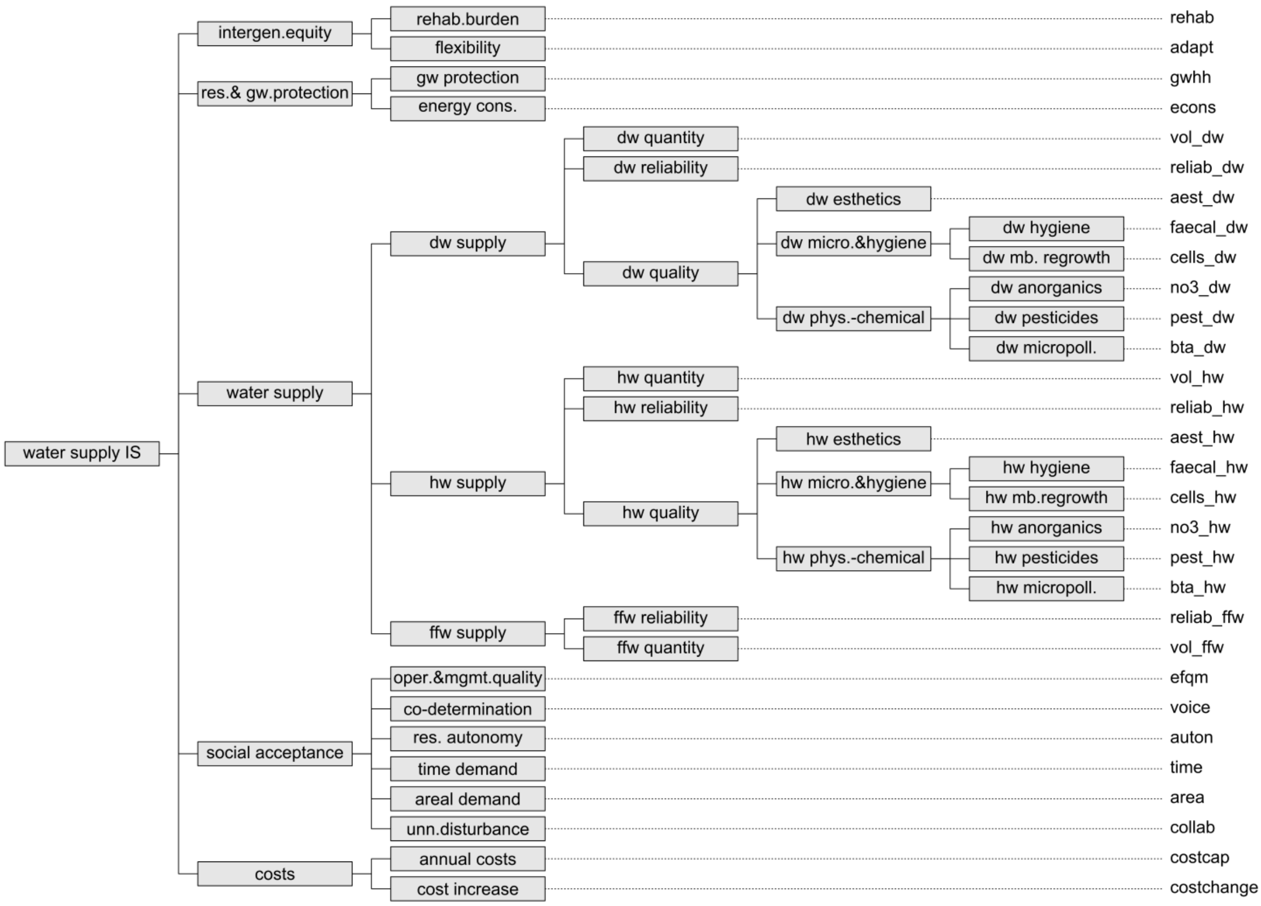


Figure 1: Objectives hierarchy for achieving the overall objective of "good water supply infrastructure". Boxes show the fundamental objectives which are connected to the corresponding sub-objectives or attributes (end of the dotted line, right edge of the plot). For more details and the meaning of the abbreviations (Table SI2.1).

Decision alternatives

Altogether eleven decision alternatives were generated in a stakeholder workshop (Lienert, et al., 2014a). 17 factors regarding organizational structure, spatial extent, technical management, and system technology were used to generate a sufficiently different set of options. Technically, these ranged from conventional centralized treatment and distribution of drinking water for all purposes (potable, household, and firefighting use) to partially or fully decentralized alternatives, e.g. with rainwater harvesting in households, in-house treatment, water delivery by lorries, or decentralized fire-fighting tanks. Different spatial extents (all or part of the utilities together, collaborations with external service providers), organizational forms (e.g. municipality-run, cooperatives, contracting), as well as technical management regimes (minimal/ moderate/

extensive inspections and maintenance; rehabilitation by condition or prioritization) were also covered. The detailed characteristics of the alternatives and their attribute predictions are compiled in the SI, section SI3.

Future scenarios

The alternatives were evaluated for four scenarios with a time horizon of 40 years (2010–2050). The future scenarios *Boom*, *Doom*, *Quality of life*, and *Status quo* (developed in a stakeholder workshop, see Lienert, et al., 2014a) cover changes in per capita income, population growth, urban expansion, water demand, and similar aspects. Hence, they define important framework conditions for the technical dimensioning of the alternatives. In the Status quo scenario, the situation in 2050 is assumed as today. There is no urbanization increase, and a stable population of ca. 24'200 inhabitants. The landscape remains rural with extensive agriculture. There is high environmental and water quality awareness. Economic growth reaches approx. 0.4 %/year of real income increase, as in the past years. The Boom scenario, despite being highly prosperous (real income increase of 4 %/year) and technologically booming, faces rapid urbanization challenges with an increased need for both densification and expansion of urban areas (200'000 inhabitants in 2050). The Quality of life scenario represents the "most desirable" scenario, with moderate, stable population and economic growth (ca. 29'000 inhabitants in 2050, real income increase +2 %/year), and high environmental awareness. In contrast, the Doom scenario represents the least desired situation with strong financial pressures (real income decrease: -1.5 %/year) and sacrifices regarding environmental protection and water quality. The urban extent, however, remains the same as in the Status quo, and the population decreases only slightly (ca. 23'000 inh. in 2050).

We judged elicitation of scenario-dependent weights or preferences as proposed by e.g. (Karvetski, Lambert, & Linkov, 2009b; Montibeller, et al., 2006) as highly hypothetic given the long time horizon. Instead, we evaluate the alternatives given current stakeholder preferences (e.g. Goodwin & Wright, 2001). This means that we use the same preferences based on which decisions are currently taken. Nonetheless, we encourage future validation and /or re-elicitation following an adaptive management approach. Even though we assume stable preferences for all scenarios, the performance of alternatives (and hence rankings) considerably differs in the scenarios, as attribute levels change under varying framework conditions.

2.2 Elicitation of preferences

For a complete MAUT analysis, the aggregation functions, marginal utility functions, and weights of all aggregation nodes (branch intersections) of the objectives hierarchy need to be elicited (e.g. Eisenführ, et al., 2010; Keeney & Raiffa, 1993). This is practically infeasible in our case, given the high complexity of the objectives hierarchy (30 marginal value and utility functions, 15 hierarchical aggregation nodes, 44 weights, Figure 1) and little elicitation time available with stakeholders. Consequently, we applied a simplified elicitation procedure followed by an uncertainty analysis. Hereby, we considered the uncertainty of not elicited components and stated stakeholder preferences.

A two-step approach combining an online survey with face-to-face elicitation was chosen. The face-to-face interview approach to preference elicitation as typical in MCDA (e.g. Eisenführ, et al., 2010; Keeney & Raiffa, 1993) and builds on the "constructive processing approach" (Hoeffler & Ariely, 1999). It assumes that preferences are constructed during the elicitation process (e.g. Belton & Stewart, 2002). Alternatively, MCDA preferences can also be formed, discussed, and elicited in groups (e.g. Kilgour & Eden, 2010; Phillips & Bana e Costa, 2007), which is often done in practice. Therefore, close interaction between the analyst and the decision maker is an important element of decision support. To avoid group effects among stakeholders obscuring individual preferences (e.g. groupthink or anticipatory consensus, see Gregory, et al., 2012; pp. 196-197), we preferred face-to-face interviews over group elicitation. The elicitation of preferences through (usually large) online surveys is more common in economics, especially when aiming at elicitation of representative "average" preferences of a population which requires large samples. They commonly rest on the assumption of existing preferences which only need to be elicited appropriately (using e.g. willingness to pay studies or discrete choice experiments; for the latter see e.g. Ben-Akiva & Lerman, 1989). This allows addressing more stakeholders in less time – a property we make use of to simplify and focus the individual face-to-face interviews.

Before the interview

All MCDA interview partners received information materials 2–6 weeks in advance, giving a short description of the purpose of the study, the decision problem, and the five top-level fundamental objectives. To avoid splitting bias (Borcherding, et al., 1991; Schuwirth, et al., 2012), the description of the five objectives was roughly equally long (299–305 words each). It contained an explanation of the objective and examples about the influence of different water supply alternatives on the achievement of the objective. The current situation was also presented. Additional material was provided, i.e. a table describing the attributes and ranges used for measurement (similar to Table SI2.1), and information about the modeling of preferences and underlying rationality assumptions (consistency, completeness, transitivity, and preferential independence). Before the interviews, stakeholders were asked to give a preliminary ranking of the objectives in an online survey to allow individual adaptation of the interviews.

Online survey

The purpose of the online survey was to rank the objectives and focus later face-to-face elicitation only on the most important ones. The objectives were ranked hierarchically, starting from the top-level, and moving downwards in the hierarchy / tree (i.e. from left to right in Figure 1). The ranges of the respective attributes were provided in a pop-up dialog as well as in separate pdf documents accessible through hyperlinks in each section. The approach used for ranking is similar to the Swing method (e.g. Eisenführ, et al., 2010) for weight elicitation, but without asking to quantify scores. Hereby, the outcomes of a hypothetical reference alternative with all objectives on the worst level were compared to the outcomes of other hypothetical alternatives having one objective each on the best level. Stakeholders then ranked these hypothetical alternatives in the order in which they preferred to improve the single objectives to their best levels. After each ranking on one hierarchical level, the stakeholders marked the objectives they judged relevant for the comparison of alternatives, and which ones could be left out ("irrelevant"). For "good household water supply", and "good firefighting water supply", stakeholders could choose to use the ranking of sub-objectives as in the drinking water case (asked first) or rank them differently in a separate step. The online survey took about 25–45 minutes to complete.

Face-to-face interviews

Three people attended each interview: the stakeholder (interviewee), the analyst (interviewer), and an assistant (taking notes, running real-time calculations to select value functions and trade-offs for later parts of the interview). It started with a reminder of the purpose of elicitation and room for questions. We emphasized that the elicited preferences are individual and subjective, and that there are no wrong answers. The elicitation took about three hours, split into three parts with 5–10-minute breaks in between.

First, the Swing weights of all 44 objectives were elicited hierarchically in a top-down manner. They were elicited as intervals (as recommended e.g. by Jessop, 2011; Mustajoki, et al., 2006), including a "best guess". The ranking of objectives from the online survey was validated in each step, before the 0–100 scores were assigned. Second, a few marginal value functions were assessed using the mid-value splitting technique, and asking for $v_{0.5}$, $v_{0.25}$, and $v_{0.75}$ (explained in Schuwirth, et al., 2012). We elicited the certainty equivalent (CE) of a 50-50 lottery between the best and worst outcomes so that marginal value functions could be converted to marginal utilities (Dyer & Sarin, 1982). Again, intervals and a best guess instead of a single value were requested. Utility independence (UI) was checked by a shortened version of the procedure described in Keeney and Raiffa (1993, pp.299-301). A 50-50 lottery which leads to either the best or the worst outcome regarding one objective was compared to the assessed certainty equivalent while the outcomes regarding the other objectives were held fixed. If the stakeholder was approximately indifferent in a situation where all remaining objectives were at their worst level, and also if they were at their best levels, it was asked if the same could be assumed for all levels in between the two extremes. If affirmed, the stakeholder was considered utility independent. The number of value functions, CE's, and UI's assessed depended on the time, but at least one value function with corresponding certainty equivalent and utility independence were elicited per stakeholder. After this, rough information about the shapes of the most important remaining marginal value functions was obtained by asking if the improvement from the worst attribute level to the mid-range was equally good as the improvement from the mid-range to the best level. This gives insight about the location of the $v_{0.5}$ value and consequently about the curvature (concave, convex, or linear). Third, consistency trade-offs were asked as in Schuwirth et al. (2012), using information from the weight and value function elicitation.

During elicitation, individual acceptance thresholds were discussed whenever a stakeholder judged attribute levels above (below) a certain threshold as unacceptable. The stakeholder specified whether the threshold should affect the overall assessment of the alternative (i.e. would it be unacceptable, no matter the level of the other attributes) or only the affected sub-objective. Thresholds were validated, and sometimes added by the stakeholders, after receiving a written summary of the elicited preference information.

2.3 Preference modeling

The stakeholders' preferences are described by individual hierarchical utility models decomposed into marginal (single-attribute) utility functions, weights, and an aggregation function, which aggregates the marginal utilities and weights to achieve one overall score for each alternative. Because the Swing method cannot be used for the weighting of marginal utility functions (as it requires the statement of preference differences, see Eisenführ, et al., 2010; p. 306), marginal values were aggregated first, and then converted to utilities on the highest level of the hierarchy. Two aggregation models were considered: the common additive aggregation model for compensatory aggregation (Keeney & Raiffa, 1993; weighted arithmetic mean) and the Cobb-Douglas model for not fully compensatory aggregation (Cobb & Douglas, 1928; weighted geometric mean). Their mathematical functions are given in Table 2. These and further aggregation models are discussed in Langhans et al. (2014).

Table **2: Hierarchical aggregation functions.** Notation: w_i weight of sub-objective belonging to value function v_i; $v = (v_1 \dots v_n)$; the weights of the additive and Cobb-Douglas model add up to unity.

Name	Function	Reference(s)
Additive model	$V(v) = \sum_{i=1}^{m} w_i v_i$	(Dyer & Sarin, 1979; Keeney & Raiffa, 1993)
Cobb-Douglas model	$V(v) = \prod_{i=1}^{m} v_i^{w_i}$	(Cobb & Douglas, 1928)

To convert multi-attribute values *V(v)* to multi-attribute utilities, the exponential model was used (e.g. Eisenführ, et al., 2010; Keeney & Raiffa, 1993):

$$U(V) = \frac{1 - e^{-rV}}{1 - e^{-r}} \tag{1}$$

Parameter *r* defines the curvature of the (overall) utility function. Unless otherwise stated, marginal values were also assumed to follow an exponential function

$$v_j(x_j) = \begin{cases} \dfrac{1 - e^{-c_j \tilde{x}_j}}{1 - e^{-c_j}}, & c_j \neq 0 \\ \tilde{x}_j, & c_j = 0 \end{cases} \tag{2}$$

where parameter c_j determines the curvature of the marginal value function $v(x_j)$ given an attribute level x_j, $x \in [x^-, x^+]$ of an attribute j and where $\tilde{x}_j = (x_j - x_j^-)/(x_j^+ - x_j^-)$.

2.4 Uncertainty analysis

Both, the prediction of the attribute levels and the preference parameters of the utility function, are uncertain in our example (as elsewhere). These uncertainties are formulated as probability distributions. From $p_a(x)$, the probability density of the attributes *x* for alternative *a*, we can compute the expected utility of an alternative *a* (Eisenführ, et al., 2010)

$$EU(a) = \int p_a(x) \cdot u(x)\, dx \tag{3}$$

Additionally, we propagated the uncertainty of preferences through the probabilistic description of the preference parameters (aggregation function, marginal value function curvature, utility function curvature, weights). This is not usually done in MCDA applications. It leads to a probability distribution of expected utilities. As the utilities have only an ordinal interpretation (Keeney & Raiffa, 1993), we calculated the resulting probability distribution of ranks of the different alternatives for each scenario.

Practically, the distribution of attribute outcomes was approximated by a random sample of n = 10'000 realizations for each alternative and scenario, assuming independence between attributes (i.e. total matrix size for four scenarios and eleven alternatives = 440'000). In a second step, we drew *s* = 10'000 times from the distribution of preference parameters and calculated the expected utilities for the attribute sample for each *s*. The expected utilities of the alternatives for each *s* were then ranked for each scenario. The distributions of the ranks were then used to compare the alternatives. We compared these results to a ranking obtained under "usual" simplification assumptions, i.e. assuming additive aggregation and linear marginal values – unless elicited in detail –, the elicited best-guess weights, and a utility function which is identical to the value function (risk neutral).

Attribute predictions

The outcomes of the attributes were predicted for all eleven alternatives and four scenarios (over 40 years; 2010–2050). Our attribute predictions stem from sources of varying quality: (1) the alternative definition (e.g. number of infrastructure sectors that collaborate in planning and construction, *collab*) or dimensioning (e.g. areal demand for water facilities in households, *area*), (2) expert estimation (e.g. aesthetic and microbial drinking water quality, *aes_dw* and *faecal_dw*; technical flexibility, *adapt*), (3) detailed models (e.g. rehabilitation demand, *rehab*; reliability of drinking water supply, *reliab_dw*), or combinations, see Table SI2.1 and Table SI3.2 for details and distributional assumptions. In the first case, the prediction of attribute levels resulted from dimensioning and no additional uncertainty was assumed. For instance, tanks were dimensioned on the maximum amount of water they need to hold, and the area demand on private property was derived from standard sizes of such tanks. In the second case, the experts' estimates (intervals) were interpreted as 90 % confidence intervals of a normal distribution, the lower value of the specified range as 5 % quantile, the upper as 95 %. From these, we obtained the mean and standard deviation. In the third case, the formulation of probability distributions was rather straightforward. Unless an appropriate distribution was known from the modeling process, an output sample was generated and different distributions (normal, lognormal, beta, gamma, logistic, truncated normal) were fitted. Using quantile-quantile and histogram plots, the best-fitting distribution was selected.

Hierarchical value (aggregation) function

Because the aggregation model was not elicited in detail, we assumed that any of the aggregation models (additive, Cobb-Douglas) or mixtures, e.g. suggested by Langhans et al. (2013; 2014), could be appropriate at each aggregation node, if preferential independence of objectives holds. Hence, the aggregation function is

$$V = \alpha_k \cdot V_{add} + (1 - \alpha_k) \cdot V_{cd} \tag{4}$$

and the mixture parameter α_k for aggregation at node k of the hierarchy was assumed to follow a uniform distribution on [0,1], where $\alpha = 1$ stands for full additivity, and $\alpha = 0$ for pure Cobb-Douglas aggregation.

Single-attribute value functions

The shape parameter of the marginal value functions was also probabilistically described, its distribution depending on the information obtained during elicitation. Three cases were distinguished:

a) **$v_{0.25}$, $v_{0.5}$, $v_{0.75}$ known:** an exponential function was fitted to the elicited intervals, assuming the uncertainty of the estimated curvature parameter c_j to follow a normal distribution $N(\mu_j,\sigma_j)$. Graphical inspection revealed that the resulting sample space when using full standard deviations of the fit was rather large and that the elicited intervals were also covered (within the 95 % confidence intervals) if only half the standard deviation was used (Figs. SI4.12a–f). Hence, the latter was done to increase specificity.

b) **Approximate shape known**: the uncertainty of the exponential curvature parameter c_j was described by a uniform distribution *Unif[min, max];* the minimum and maximum were chosen as follows: *Unif[0,10]* if concave, *Unif[-10, 0]* if convex, and *Unif[-0.4,0.4]* if approximately linear.

c) **No information:** exponential function with $c_j \sim$ *Unif[-10,10]*

Hierarchical utility function

Since we did not elicit the aggregation parameters for utilities but only the parameters to aggregate values (section 2.3), we aggregated values up to the highest hierarchy level. The aggregate overall value was then converted into an aggregate overall utility assuming an exponential function with $r \sim$ *Unif[-10,10] (Eq.1).*

Weights

The elicited "best guess" weight and intervals (equally spaced around best guess) were interpreted as centered 95 % confidence intervals and probabilistically described as normal distributions truncated at [0,1]. The mean value μ_i was the best guess and the stated intervals were interpreted as $\pm$ 1.96 times the standard deviation σ_i to cover the 95 % interval (SI, Table SI4.1). Weights were then independently sampled within each (sub-) branch of the objectives hierarchy and normalized to 1 (dividing by their sum), as required by the additive and Cobb-Douglas model. Therefore the resulting weights no longer follow a truncated normal distribution. Other sampling techniques and their implications are described in Butler et al. (1997) and Mustajoki (2011).

Acceptance thresholds and individual adjustments

Acceptance thresholds were implemented as external elimination criteria by setting the overall value and utility of an alternative or branch (depending on the specification by the stakeholder) to zero if the predicted attribute level exceeded (or fell below) the threshold. Some of the stated thresholds for the "days per year with hygienic concerns of drinking water" were stricter than the estimated attribute level of the current system. For example, some stakeholders set the threshold to 0 or 2 days per year while the status quo (estimated by an expert) lies between 0 – 5 days per year. Current legal guidelines require that no fecal indicator bacteria are found during microbial screenings. This might have motivated the respective stakeholders to set such extremely low acceptance thresholds (see SI, Table SI4.6). Microbial screening is done approx. 1–6 times per year depending on the water supplier and fecal bacteria have been (rarely) detected, leading the expert to estimate that up to five days of water quality impairment per year are currently possible. To reconcile this, thresholds were adjusted to allow for the status quo of max. 5 d/a, implying that stakeholders find the current situation acceptable. Additionally, the exponential distribution is not steep enough to cover the stated intervals of two stakeholders over the whole attribute range from 0–365 d/a (SH5, 6). Therefore, the function was estimated on a range from 0–30 d/a, also in line with the acceptance threshold of SH5, and assumed as zero for higher attribute levels (SI, Figure SI4.12e). SH10 specified a step function for the attribute "% utilization of groundwater recharge" with absolute certainty (SI, Figure SI4.12b), which we used instead of an exponential function.

2.5 Global sensitivity analysis (GSA)

The magnitude of influence of the preference parameters on the alternatives' rankings was calculated with a variance-based global sensitivity analysis, following the extended Fourier Amplitude Sensitivity Test (e-FAST) approach by Saltelli et al. (Saltelli, 2008; Saltelli, et al., 2006; A. Saltelli, S. Tarantola, & K. P. S. Chan, 1999b). An application to a simple MAVT-problem is presented in Saltelli et al. (1999a). We considered 90 uncertain parameters $\boldsymbol{\theta}$, including 44 weights w_i, 30 marginal value function curvatures c_j, 15 mixing parameters a_k for aggregation, and one utility function curvature r.

The first and total order coefficients of the preference parameters (i.e. parameters of the hierarchical utility function) were calculated. The first order sensitivity coefficient S_z measures the main (individual) effect of the parameters $\boldsymbol{\theta}=\theta_1\ldots\theta_z$ (Saltelli, et al., 2010):

$$S_z = \frac{Var_{\theta_z}(E_{\boldsymbol{\Theta}_{-z}}(Y|\theta_z))}{Var(Y)} \tag{5}$$

$Var(Y)$ stands for the variance of the model output Y, $E_{\theta_{-z}}(Y|\theta_z)$ for the conditional expectation (mean) of Y, if all parameters $\boldsymbol{\theta}$ are allowed to vary except θ_z, and $\boldsymbol{\theta_{-z}}$ stands for all parameters except θ_z. The total order coefficients S_{Tz} measure the interactive effect of changes of individual parameters with other parameters,

$$S_{Tz} = \frac{E_{\boldsymbol{\Theta}_{-z}}\left(Var_{\theta_z}(Y|\boldsymbol{\Theta}_{-z})\right)}{Var(Y)} = 1 - \frac{Var_{\boldsymbol{\Theta}_{-z}}\left(E_{\theta_z}(Y|\boldsymbol{\Theta}_{-z})\right)}{Var(Y)} \tag{6}$$

and $Var_{\boldsymbol{\Theta}_{-z}}\left(E_{\theta_z}(Y|\boldsymbol{\Theta}_{-z})\right)$ represents the first order effect of $\boldsymbol{\theta_{-z}}$.

The uncertain model output Y was represented by the Kendall correlation coefficient $\tau \in [-1,1]$ between the ranking for given parameter values $\boldsymbol{\theta}$ and the standard ranking. (Kendall, 1938). Kendall-τ is a commonly used statistic to measure the relationship between two rankings. τ equals 1 if the compared rankings are identical, -1 if they are completely opposite, and 0 if there is no relationship. Here, the rankings resulting from parameter changes were compared to a reference ranking (obtained using the mean preference parameters). Thus, for each sample from the joint distribution of preference parameters $\boldsymbol{\theta}$ (section 2.3), the expected utility of the eleven alternatives was calculated. The attribute distributions were represented by a discrete, independent sample with sample size reduced to n=1'000, because the rankings were nearly identical to those obtained from the larger n=10'000 sample (Figure SI6.1). The necessary parameter sample size s to achieve approximately stable sensitivity coefficients was iteratively determined. The preferences of SH2 (local operational personnel) were used as a "base case". Although termed *global* sensitivity analysis, the sensitivity coefficients depend on the distributional assumptions regarding the uncertain parameters in the 90-dimensional parameter space. Therefore, we defined five analytic GSA layouts to address specific research questions (Table 3).

Table 3: Five analytic layouts for global sensitivity analysis. SH = stakeholder.

Layout	Assumptions	Research question
SH2_SQ ("base case")	Preferences of SH2, same parameter assumptions as for uncertainty analysis. Attribute predictions for Status quo scenario.	Which are the parameters that the results are most sensitive to and which elicitation should be focused on, given the current layout for a specific stakeholder SH2?
SH2_SQ_red	As SH2_SQ, but with reduced range of value and utility function curvature parameters c_i and r, ranging from -5 to 5.	What is the effect of the size of the selected parameter sampling region on the sensitivity of parameters for stakeholder SH2?
SH2_SQ_noAT	As SH2_SQ, but without external acceptance thresholds.	How sensitive are the results to individual parameters if no external acceptance thresholds are considered?
SH2_BO	Preferences of SH2, same parameter assumptions as for uncertainty analysis. Attribute predictions for Boom scenario.	Are the same preference parameters the most influential both in the Status quo and the highly dynamic Boom scenario?
NoPref_SQ	No preferences elicited; $0< w_i <1$ (uniform); $-10< c_i <10$ (uniform); $0< a <1$ (uniform); $-10< r <10$ (uniform). Attribute predictions for Status quo scenario.	If no preferences are known, which parameters are the results most sensitive to?

2.6 Implementation

Most of the preference and uncertainty modeling was implemented in R (R Development Core Team, 2011). The R package *utility* (Reichert, Schuwirth, & Langhans, 2013) was used to implement and evaluate the MAUT model. We used the following packages for parameter optimization, estimation of the underlying failure model parameters, global sensitivity analysis and visualization: *optimx* (Nash & Varadhan, 2011), *DEoptim* (Mullen, Ardia, Gil, Windover, & Cline, 2011), *adaptMCMC* (Scheidegger, 2012), *sensitivity* (Pujol, Iooss, & Janon, 2012; assuming M = 4), and *ggplot2* and *reshape* (Wickham, 2007, 2009). The online survey was set up in a trial version of *Qualtrics* (Qualtrics, 2012).

3 Results of the case study

3.1 Attribute outcomes

Some attribute outcomes in the Boom scenario differ substantially from those of the other three scenarios (Table SI3.2, Figs. SI3.1a-f). This is the case for attributes whose performance is strongly linked to the scenario assumptions (section 2.1, "Future scenarios"): "realization of the rehabilitation demand" (*rehab*), "utilization of groundwater resources" (*gwhh*), "system reliability" (of drinking, household, and firefighting water; *reliab_dw/ hw/ ffw*), "changes in total cell counts" (drinking and household water; *cells_dw/hw*), "hygienic concerns" of drinking water (*faecal_dw*), "available water for firefighting" (*vol_ffw*), "annual cost in % of mean taxable income" (*costcap*), and "mean annual cost increase (*costchange*)". Besides their impact on the ranking of alternatives, these attributes furthermore discriminate between rankings in the four scenarios, i.e. allow to assess the stability given different boundary conditions.

Other outcomes do not differ or change only slightly as a result of the scenario assumptions. This is true for all attributes linked to "high social acceptance" (*efqm, voice, auton, time, area, collab*), but also for "energy consumption for water treatment and transport" (*econs*), "flexibility of technical extension or deconstruction of infrastructure" (*adapt*), "days per year with esthetic impairment" (of drinking/ household water; *aes_dw, aes_hw*), and "hygienic concerns" for household water (*faecal_hw*). The respective ranking of alternatives concerning these attributes is thus robust in all scenarios.

Finally, the predicted levels of some other attributes are identical and hence do not help to discriminate alternatives, but could be important in other cases and were thus not removed. In the case of "water quantity limitations" of drinking water (*vol_dw*), the outcome is zero days per year for all alternatives and scenarios, its evaluation could thus be discarded. Similarly, absence of detailed predictions for the attributes of the "high physico-chemical quality" of drinking and household water (*no3_dw/ hw, pest_dw/ hw, pest_bta/ hw*) does not support better differentiation of alternatives, but adds uncertainty (the overall attribute ranges were assumed). Whether differences in the predictions for these attributes have an impact and efforts should be spent on reducing this uncertainty, cannot be concluded without a more detailed sensitivity analysis covering also the uncertainty of the attribute parameter predictions. Details regarding individual alternatives are discussed in section 3.3 where appropriate.

3.2 Stakeholder preferences

Weights

The top-level objective ***"good water supply"*** (Table SI4.1, and Figure SI4.1) received the highest weights, scoring between 0.23–0.39 for stakeholders (SH) 1–9, and 0.35–0.43 for SH10 (second place, overlaps with 'resources and groundwater protection'). Of its sub-objectives (Figure SI4.2), "good drinking water supply" was the most important for nine of ten stakeholders and second for SH5. "Good household water supply" was eight times second (third for SH4, SH6), ranging from 0.28–0.83. Consequently, "good firefighting water supply" was eight times in the third place (second for SH4, SH6). The sub-objective ***"high social acceptance"*** was considered least important by all stakeholders with weights between 0–0.15, except SH9, who rated the weight between 0.17–0.23 (third). Two stakeholders (SH1, SH10) would even discard "high social acceptance" and four others (SH4-6, SH8) assigned zero weight to some of its sub-objectives (Figure SI4.11).

The ranks and weights of the remaining three top-level objectives were more divergent. Due to the weight variations of these objectives, the ranking of the alternatives may substantially differ depending on the stakeholder. There are no clearly visible grouping patterns of stakeholders based on the weight information alone. The ranking of objectives in the online surveys was very similar to that of the face-to-face interviews, but the judgment about the relevance of objectives was not (Table SI4.3). Stakeholders marked considerably more objectives as "irrelevant" if asked online (ca. 40 %), than during face-to-face interviews (ca. 10 %).

Marginal value functions

We obtained preference information for 172 value functions (sum over all stakeholders). The shape was elicited in detail for 21 of them (Figs. SI4.12a–f; summary Table SI4.4). Non-linear shapes were most frequent (88 concave, 61 convex), 23 functions were linear. The shape of the marginal value functions differed between and within stakeholders and objectives (e.g. SH4, SH8).

Certainty equivalents

As shown from the fitted marginal utility function parameter *r*, half of the stakeholders were intrinsically risk averse for specific objectives (10 out of 21), and about a quarter risk prone (6) or risk neutral (5), Table SI4.5). The direction (not the magnitude) of the risk attitude across several objectives was identical for three stakeholders (SH2, SH3, SH5; risk averse) and differed conditional on the objective for four others (SH1, SH4, SH8, SH10). For the remaining three stakeholders one marginal certainty equivalent was elicited for each.

Acceptance thresholds

Eight stakeholders specified acceptance thresholds (AT) that need to be considered when evaluating the alternatives (Table SI4.6). They concern either specific attribute levels, or a perceived loss / deterioration regarding some of the attributes compared to the current situation. They most commonly addressed drinking water quality concerns, specifically "days per year with hygienic concerns" (mentioned by eight of ten stakeholders, AT's at 0 d/a, 2 d/a, or 30 d/a). Others concerned the amount of groundwater abstraction, the cost increase, or the reliability of the firefighting water system. In all cases, the overall value of the alternative is affected (set to zero if AT is exceeded) and not only the value of the sub-objective.

Utility independence conditions

Six stakeholders stated that their certainty equivalent might change slightly, if the levels of the remaining attributes were extreme (on the best or worst level; see Table SI2.1). The stakeholders did not however, relate this to other individual attributes, but only to the overall performance of the alternative. We are not aware of similar observations in other studies or of behavioral effects that could explain this behavior. Arguably, this could be considered a deviation from theory. Since the stakeholders agreed that this deviation was small and could be remediated by slightly increasing the stated interval of the certainty equivalents, we were willing to accept that this would suffice to reconcile stakeholder preferences with the assumption of utility independence.

3.3 Ranking of alternatives and uncertainty analysis

To find out whether there are alternatives which are clearly best for all stakeholders or can be suggested as potential compromise, we first present the rankings of alternatives for all stakeholders and future scenarios, before looking into rankings for individual stakeholders. The differences in the rank distributions considering uncertainty are explained with help of the median rank (MR) and inter-quartile ranges (IQR) (Figure 2; Table SI5.2). An alternative is better if its median rank is smaller (i.e. approaching the first rank), and for overall risk-averse stakeholders presumably also if its IQR is narrower (e.g. if several alternatives have the same median).

There is no single alternative which is clearly best or worst for all stakeholders and all scenarios. Calculating the average over the median (or mean) ranks of the ten stakeholders (Table A1, Appendix), alternative A6

("Maximal collaboration, centralized") is best in the Doom and Quality of life scenario and second after A1b ("Centralized IKA") in the Status quo scenario. In the Boom scenario, A2 ("Centralized IKA, rainwater stored") is the best alternative, followed by A1b which also performs well in the Doom and Quality of life scenarios. Among the worst alternatives are A3 ("Fully decentralized") and A5 ("Decaying centralized infrastructure, decentralized outskirts"), besides A9 ("Centralized, privatization, minimal maintenance"; Doom, Quality of life scenarios only). Most stakeholders could be manually classified into two groups by their rank distribution patterns. They are marked either with a circle or a triangle (Figure 2).

Ranking of alternatives considering uncertainty – circle group

The *"circle group"* consists of SH2, SH4, and SH9. Its rank distribution pattern varies between the Boom and the other scenarios (Figure 2). For the Doom, Quality of life, and Status quo scenarios, alternative A6 *("maximal collaboration, centralized"*, orange lines) is the best, with a median rank of 1–2 and little overlaps (small IQR: 1–5) for all three stakeholders. The ranking of the other alternatives is less clear in the Boom scenario and differs by stakeholder. Two alternatives perform similarly in the Boom scenario, A2 (*"Centralized IKA"*, grey, MR 1–2, IQR 1–3; best for SH2 and 4) and A1b (*"Centralized, IKA, rain stored"*, lower red, MR 1–4, IQR 1–4; best for SH9). The worst alternative is A9 ("*Centralized, privatization, minimal maintenance*", pink, MR= 11, IQR= 8–11), for all scenarios. This low ranking of A9 can in part be explained by the acceptance thresholds (5 d/a) and comparatively high weights concerning the attribute "d/a with hygienic concerns of drinking water quality" (*faecal_dw*, section 2.4, "Acceptance thresholds and individual adjustment"; and Figure SI4.5; Table SI4.6). According to the attribute predictions, those of A9 could sometimes exceed threshold of 0–10 d/a, thus explaining its low performance (Fig SI3.1b). For Alternative A6 this attribute was predicted to be 0 d/a for all scenarios. Consequently it was not penalized by this threshold. It also performs well regarding a range of other highly-weighed attributes, namely "realization of the rehabilitation demand" (*rehab*, ca. 25–80 %, Figure SI3.1a), "flexibility of technical extension or deconstruction of infrastructure" (*adapt*, ca.45–65 %, Figure SI3.1a), "system reliability" of drinking, household, and firefighting water (*reliab_dw/ hw/ ffw*, all <0.01, Figure SI3.1b/c/e), "d/a with esthetic impairment" of drinking water (*aes_dw*, 0–10 d/a, Figure SI3.1b), and the cost attributes (*costcap, costchange*, ca. 0.01 % of mean income and < 0.8 % increase per year, Figure SI3.1f). Its poor performance concerning "utilization of groundwater recharge" (*gwhh*, ca. 80–150 %, Figure SI3.1a) in the Boom scenario explains why it is not one of the best alternatives in that case. In the Boom scenario, A1b, and A2 perform similarly well regarding most of these objectives, and outperform A6 regarding "utilization of groundwater recharge" (*gwhh*, <15 %), and some of the household water quality attributes (*aes_hw, faecal_hw, cells_hw*, Figure SI3.1b-d).

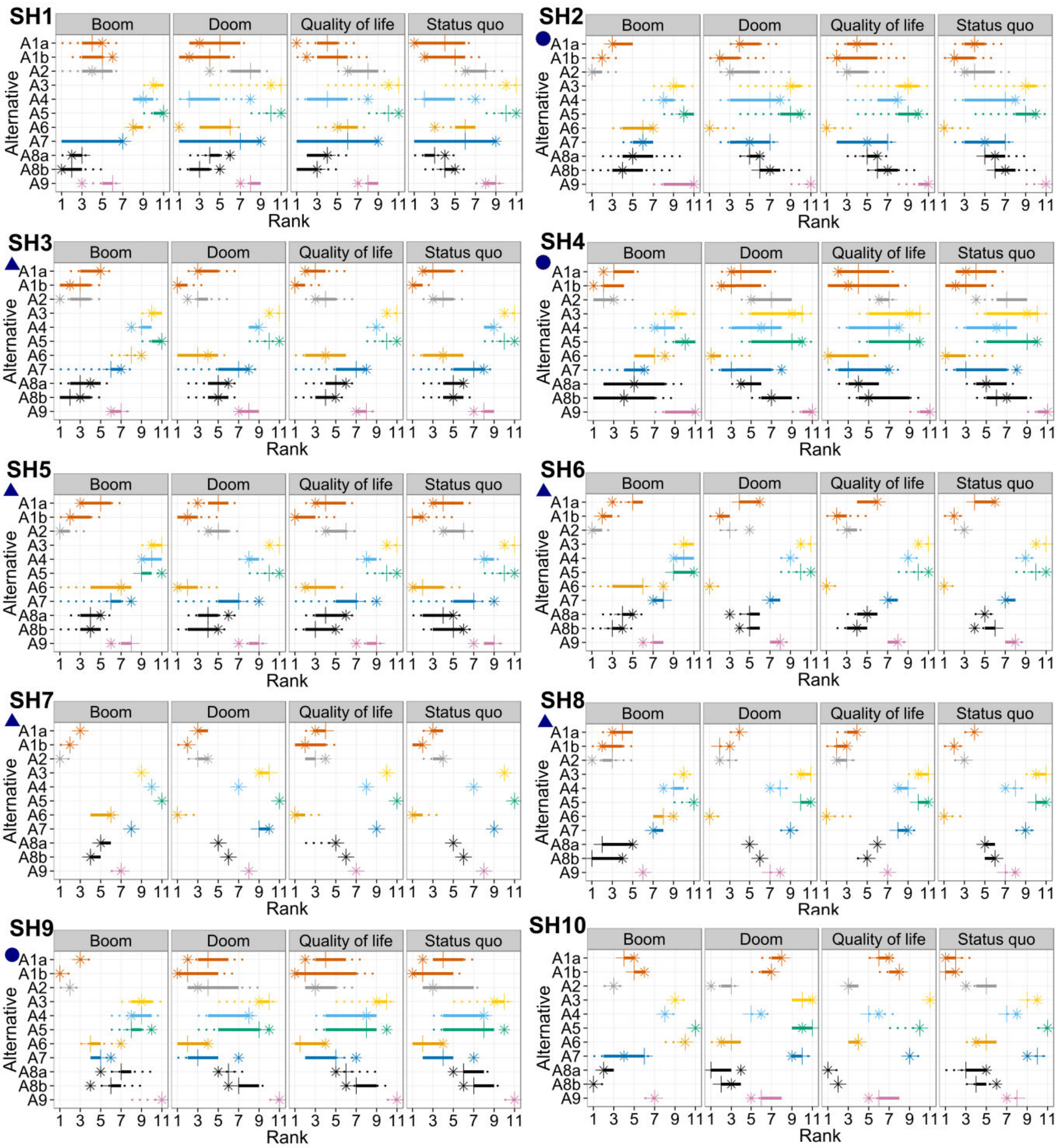


Figure 2: Ranking of eleven alternatives for water supply infrastructure for all ten stakeholders and four future scenarios considering uncertainty. "1" indicates the best rank and "11" the worst. Thick, solid lines with vertical dash represent the 25 %, 75 %, and 50 % quantiles (lower quartile, upper quartile, median), dotted lines the 5 % and 95 % quantiles. Stars indicate the expected ranking with resulting from the "usual" simplification assumptions (section 2.4). Some stakeholders can be classified into two distinct groups, indicated with a filled circle or triangle (top-left in each figure).

Ranking of alternatives considering uncertainty – triangle group

The "*triangle group*" has five members, SH5–8, and SH3 (Figure 2). Like in the *circle group*, alternative A6 (orange) clearly has the lowest median rank (1–3) and smallest IQR (1–2) for SH6-8, but not for SH3 and SH5 (A1b best regarding MR and IQR, blue) in all scenarios except Boom. The worst alternatives in the Doom, Quality of life, and Status quo scenarios for all group members are either A3 ("Fully decentralized", yellow) or A5 ("Decaying infrastructure everywhere", green); both with mean rank 10–11 and IQR 9–11. In the Boom scenario, the individual best and worst ranks are less clear. Better-ranked alternatives are: A2 (MR 1–3; best for SH5-7) and A8b ("Status quo, technical variant", MR 1–4; best for SH3 and SH8). The worst alternative is either A3 (MR 9–11, IQR 9–11), A4 (MR 9–10, IQR 9–11), or A5 (MR 9–11, IQR 9–11). Thus, the performance of alternatives in the triangle group cannot be easily explained by one single preference characteristic. None of the stakeholders stated strict acceptance thresholds (Table SI4.6), which might influence the ranking.

Ranking of alternatives considering uncertainty – SH1 and SH10

The ranking for SH1 is similar to the triangle group, but especially affected by attributes with diverging predictions in the four scenarios. The best alternative with the lowest median rank (= high expected utility) is either A8b (Boom, Quality of life), or A4 (Doom, Status quo). This is linked to the outcomes of A4 for "system reliability" of drinking, household, and firefighting water (*reliab_dw/ hw/ ffw*). They perform well in the Doom and Status quo scenarios, because the amount of decentralized assets in A4 is small, leading to a higher reliability of the system, and ultimately to A4 being the best-performing option. The reliability is lower in the other scenarios, because the proportion of decentralized assets increases, and A8b is best instead (Figure SI3.1b–e). The performance regarding costs might also have an impact in the Boom scenario, since the weight of costs is comparatively high for SH1 (see Figure 2). The elicited value function for annual cost increases (*costchange*) is strongly convex for SH1 which leads to a stronger decreasing marginal value in the case of high cost increases compared to low increases as expected with A8b (Figure SI4.12f). The worst alternative for SH1 is either A3 or A5, and the same reasoning applies as for the triangle group. Additionally, A3 and A5 might lead to high cost increases, further penalizing their outcome.

Regarding SH10, A8b has the best median rank of 1 in the Boom and Quality of life scenario. In the Doom scenario, A8a would be best (MR: 1, IQR: 1–3), and in the Status quo A1b (MR 1-2). The worst alternative is either A3 (MR 9–11, IQR 9–11) or A5 (MR 10–11, IQR 9–11). As SH10 discarded "intergenerational equity" and "social acceptance", the remaining top- and lower-level objectives have comparatively high weights, such as the "natural groundwater balance" or "high supply reliability" (Table SI4.1). Additionally, the value function over the corresponding attribute ("% utilization of groundwater recharge") for this objective is unity whenever 100 % or less of groundwater recharge are abstracted (Figure SI4.12b), otherwise the whole alternative is unacceptable (see ATs, Table SI4.16). This explains the poor ranking of A6 in the Boom scenario (likely exceeding 100 %; Figure SI3.1a). The ranks of A3 and A5 appear in line with the predictions for drinking, household, and firefighting water reliability. Compared to others, these are very high in A3 and A5; Figure SI3.1b-c, Figure SI3.1f).

Simplifying assumptions

In some cases, the ranking obtained under usual simplifying assumptions (linear marginal value functions unless elicited in detail, additive aggregation, best-guess weights, risk neutrality) deviates considerably from the rankings obtained with the uncertain parameters. This is indicated by the divergence between the stars (usual simplifying assumptions) and vertical dash (median rank) in Figure 2. For example, in the Doom scenario, alternative A6 would receive a much better rank of 1 for stakeholder SH1 under usual simplification assumptions, while its median rank is only 6. Opposed to this, the ranking of alternative A4 in the Doom, Quality of life, and Status quo scenario would be clearly inferior (usual simplification assumptions: rank 8, 8 and 7, MR: 2, 4, 2). Differences between the rankings under usual simplification assumptions and uncertain preferences are most frequent for SH1, SH4, SH5, and SH9, but do not substantially depend on specific alternatives or scenarios (see also Table SI5.3). Despite these individual differences, the mean rank across the ten stakeholders would lead to identify the same candidate best alternatives (or: potential compromise solutions), as indicated by the mean of the individual median ranks used above (Table A1 , Appendix).

4 Results of the global sensitivity analysis

The first (S_z) and total order (S_{Tz}) sensitivity coefficients for the five analysis layouts (cf. Table 4) were calculated with s= 4'000, i.e. a sample matrix of 360'000 rows by 90 columns. With this sample size, the first order coefficients S_z were approximately stable, but not the total order coefficients S_{Tz} of which only the rank was approximately stable (Table SI6.1, Figure SI.6.2). Increasing s further was not feasible due to the computational expense (with s = 4'000 implemented as 72 parallel runs per GSA layout, each ran about 3 days on a high-performance computation cluster). Nevertheless, we judged knowledge of the first order coefficients and a ranking of total order coefficients sufficient for the interpretation of the results. Thus, first order coefficients are interpreted quantitatively and total order coefficients qualitatively, based on their ranking.

The sum of the S_z is below unity for all five model layouts, see Table 4 ($\sum\theta_z$). This means that the models are considerably nonlinear and that 30–77 % of the output variance (of the rank correlation coefficient τ) can be explained by the variation of individual parameters alone (main effect). In the base case (SH2_SQ), 77 % of the output variance is explained solely by the main effects, which can be attributed by a large extent to the overall utility function curvature r, accounting for ca. 72 % of the output variance. This parameter also has the largest total order coefficient (rank 1, see also Figure 4), demonstrating high interactive effects with other parameters. If the top-five ranking parameters were known with certainty (*r, a.IE, a.overall, c.IE_rehab, a.SA*; i.e. the overall risk attitude, aggregation mixture parameter of "high intergenerational equity", overall aggregation mixture parameter of "good water supply infrastructure", marginal value function curvature of "low rehabilitation demand", aggregation mixture parameter of "high social acceptance"), more than 75 % of the uncertainty of the ranking could be resolved (Table 4; see "∑Rank 1–5"). All other parameters are much less relevant, with main effects < 2 % and considerably lower interaction effects (smaller red bars in Figure 4). The utility function curvature r is also clearly the most important parameter regarding its affect on output uncertainty when the Boom scenario (SH2_BO) or reduced parameter ranges (SH2_SQ_red) are considered. In that case, r explains 25 % and 31 % of the overall variance by its main effect, respectively, and is also the most important parameter regarding interactions. However, uncertainty about other parameters such as the aggregation mixture parameters ratios α_k and marginal value function curvatures c_j becomes more influential, visible in the top-ranked parameters and the sums of the respective grouped sensitivity coefficients (Table 4, lower part).

The high importance of the utility curvature parameter r can be explained by the distributions of the values, see Figure 3. Due to the external acceptance thresholds (ATs), some alternatives (A1a, A1b, A2, A8a, A8b, A9) have extremely wide overall value distributions, reaching from zero to values above 0.85. If SH2 is risk averse (r>0), alternative A6 will always perform best in the Status quo scenario, because it is considerably less uncertain. Consequently, the ranking of A8a/b compared to A7 is affected by the risk attitude, if ATs apply.

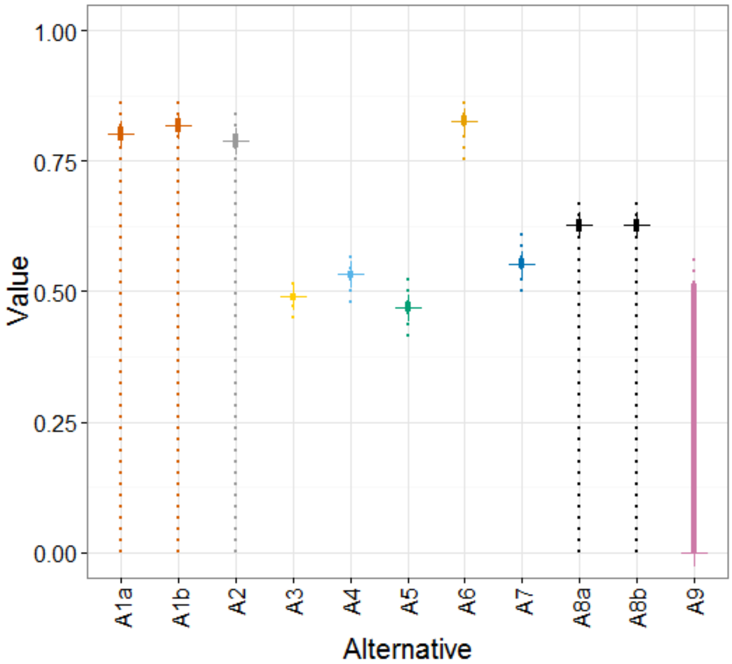


Figure 3: Overall value of the alternatives using the mean preference parameters for SH2 in the Status quo scenario. The 50 % quantile is represented by horizontal bars, the upper and lower quartile by the solid vertical line and the extremes by the dotted vertical lines.

If the ATs are not accounted for, uncertainties of other parameters become more influential, reflected in the respective sensitivity coefficients ("SH2_SQ_noAT", Table 4). Not only is the non-linearity of the model considerably higher (main effect explains only 39 % of output variance, 61 % due to interactions), but also the sensitivity to towards the risk attitude is much lower (negligible main effect, low interactive effect; see also Figure 4 "SH2_SQ_noAT"). If ATs are not considered, the marginal value function curvatures c_j are the most influential parameters (grouped main effect: 31 %, eight out of fifteen top ranked parameters).

If no stakeholder preferences are considered (NoPref_SQ), about 70 % of the output uncertainty arises from interactions between uncertain parameters, and only 30 % can be explained by individual effects. Ten out of the fifteen most influential parameters by individual effect are weights. Also, the summarized main effect of the weights group ($\sum w_i$ = 17 %, considerable interactions, see also Figure 4, NoPref_SQ) suggests that the model is most sensitive to the weights – quite contrary to their lower sensitivity in the four other cases. Once again, the sensitivity to the overall utility curvature parameter r is negligible in the absence of ATs.

Table 4: Sensitivity coefficients of five analysis layouts. The assumptions underlying the five layouts are summarized in Table 3. Only the top 15 parameters with highest first order effect (S_z, upper part of the table) and sums over parameter groups (lower part of the table) are shown. Parameter *r* is the overall risk attitude, parameters starting with "*a.*" the aggregation mixture parameters, "*c.*" value function curvature parameters, and "*w.*" the weighting parameters. Parameter names begin with the parameter group ("*a.*" or "*c.*"), followed by the respective main objective of the branches going down the hierarchy up to the indicated end point (aggregation node or attribute, see Figure 1). Acronyms for the top-level main objectives are: "IE" – *high intergenerational equity (w.1)*, "RG" – *high resources and groundwater protection (w.2)*, "WS" – *good water supply (w.3)*, "SA" – *high social acceptance (w.4)*, and "KO" – *low costs (w.5)*. E.g. "*c.WS_dw.reliab*" stands for the value function curvature of the objective *high reliability (reliab) of the drinking water supply (WS_dw)*. "*a.overall*" – is the mixture parameter at the highest hierarchical level. The weight numbers are given in Table SI4.1.

	SH2_SQ ("base case")			**SH2_SQ_red**			**SH2_SQ_noAT**			**SH2_BO**			**NoPref_SQ**		
Rank	**θ_z**	**S_z**	**Rank (S_{Tz})**	**θ_z**	**S_z**	**Rank (S_{Tz})**	**θ_z**	**S_z**	**Rank (S_{Tz})**	**θ_z**	**S_z**	**Rank (S_{Tz})**	**θ_z**	**S_z**	**Rank (S_{Tz})**
1	***r***	**0.717**	**1**	***r***	**0.308**	1	*c.RG_energ*	0.111	2	***r***	**0.248**	1	*a.IE*	0.117	1
2	*a.IE*	0.019	3	*c.IE_rehab*	0.090	2	*c.IE_flex*	0.088	1	*c.IE_rehab*	0.100	3	*w.1*	0.098	2
3	*a.overall*	0.010	2	*a.IE*	0.068	3	*c.SA_area*	0.041	7	*c.RG_gwhh*	0.055	2	*w.3*	0.044	3
4	*c.IE_rehab*	0.010	5	*a.SA*	0.015	6	*c.SA_auton*	0.031	8	*a.IE*	0.026	6	*w.1.2*	0.007	4
5	*a.SA*	0.003	6	*c.IE_flex*	0.007	5	*a.SA*	0.018	6	*c.SA_collab*	0.010	8	*w.4.2*	0.004	7
6	*a.WS_dw*	0.002	8	*a.overall*	0.007	4	*c.IE_rehab*	0.018	4	*a.SA*	0.009	7	*c.IE_rehab*	0.003	63
7	*c.IE_flex*	0.002	4	*a.WS_dw*	0.003	7	*a.IE*	0.011	5	*a.overall*	0.005	4	*a.SA*	0.002	5
8	*c.WS_dw.reliab*	0.001	10	*c.WS_dw.reliab*	0.003	29	*a.overall*	0.008	3	*c.SA_efqm*	0.004	38	*c.RG_energ*	0.002	19
9	*c.WS_ffw.quant*	0.001	17	*c.SA_collab*	0.002	10	*a.WS_hw*	0.008	15	*c.SA_auton*	0.003	11	*w.4.6*	0.002	9
10	*c.RG_energ*	0.001	7	*c.RG_energ*	0.002	9	*c.SA_efqm*	0.008	14	*c.IE_flex*	0.003	5	*w.4.1*	0.002	20
11	*c.SA_time*	0.001	23	*c.SA_time*	0.001	14	*a.RG*	0.005	11	*c.RG_energ*	0.002	10	*w.2.2*	0.002	8
12	*c.SA_auton*	0.001	16	*c.WS_ffw.reliab*	0.001	17	*a.WS_dw*	0.003	9	*a.WS_dw*	0.002	9	*a.overall*	0.002	6
13	*c.SA_efqm*	0.001	11	*a.WS_hw*	0.001	56	*c.RG_gwhh*	0.003	65	*c.WS_dw.reliab*	0.002	47	*w.4.4*	0.002	16
14	*c.WS_ffw.reliab*	0.000	12	*a.RG*	0.001	57	*w.4.5*	0.002	12	*c.SA_area*	0.001	37	*w.3.2.2*	0.001	25
15	*w.2*	0.000	36	*c.WS_w.reliab*	0.001	20	*c.SA_collab*	0.002	13	*a.WS_hw*	0.001	62	*w.3.3.1*	0.001	17
	$\sum\theta_z$	**0.771**		***$\sum\theta_z$***	**0.524**		***$\sum\theta_z$***	**0.386**		***$\sum\theta_z$***	**0.483**		***$\sum\theta_z$***	**0.297**	
	∑Rank 1-5	**0.758**		***∑Rank 1-5***	**0.488**		***∑Rank 1-5***	**0.288**		***∑Rank 1-5***	**0.438**		***∑Rank 1-5***	**0.269**	
	$\sum w_i$	0.002		***$\sum w_i$***	0.007		***$\sum w_i$***	0.017		***$\sum w_i$***	0.006		***$\sum w_i$***	0.166	
	$\sum c_j$	0.017		***$\sum c_j$***	0.114		***$\sum c_j$***	0.312		***$\sum c_j$***	0.185		***$\sum c_j$***	0.009	
	$\sum \alpha_k$	0.035		***$\sum \alpha_k$***	0.095		***$\sum \alpha_k$***	0.055		***$\sum \alpha_k$***	0.045		***$\sum \alpha_k$***	0.122	

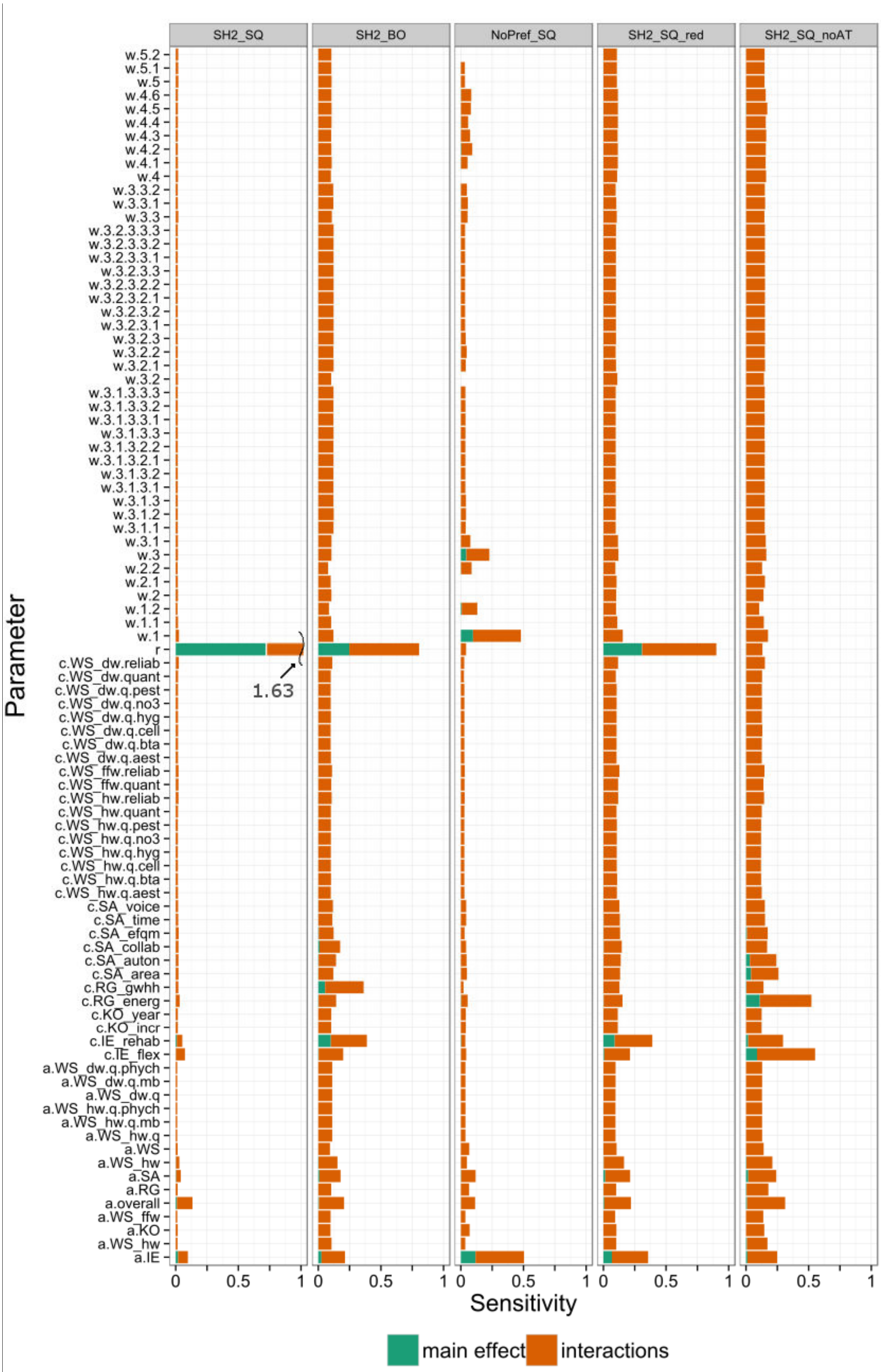


Figure 4: First (S_z) and total order (S_{Tz}) sensitivity coefficients for the analytic layouts presented in Table 3. The S_z (“main effect”) are represented by green bars, S_{Tz} (“interactions”) by red bars. Description s. Table 4.

5 Discussion

5.1 Water infrastructure planning in the case study

Regarding the "Mönchaltorfer Aa" case study, our ten stakeholders were unanimous about "good water supply" being of highest and "social acceptance" of lowest priority for achieving the overall goal of "good water supply infrastructure" (given the assumed attribute ranges). The relative importance of the other objectives diverged more strongly between stakeholders.

No single best or worst alternative for all stakeholders or scenarios could be identified, as the individual ranking distributions differed and sometimes overlapped (section 3.3). Nonetheless, suggestions for potential compromise solutions can be made: Under the presented assumptions, alternative A6 ("Maximal collaboration, centralized") might be a good compromise since it performed well for many of the highest-weighted objectives, namely water supply reliability and drinking water quality, intergenerational equity (technical flexibility, rehabilitation demand), but also for costs. A6 consequently had the best mean rank in the Doom and Quality of life scenario and the second rank in the Status quo scenario. In the Boom scenario, however, it performed considerably worse than the best-ranked alternative A2 ("Centralized IKA, rainwater stored", Table SI3.1), due to its high "utilization of groundwater recharge", compared to A2 and other well-performing alternatives. A2 and A6 are technically similar, also to the poorly performing A9 ("Centralized, privatization, minimal maintenance"). They are basically adaptations of the current centralized supply system, but all foresee pipe dimensioning for residential areas based on household peak demands, rather than the (higher) firefighting peak demands. Firefighting is accounted for by decentralized, shared firefighting tanks. Another (relevant) difference is the poor management and rehabilitation in A9, compared to moderate efforts undertaken in A2 and A6. Also the limitation of water imports from external sources in A6 to max. 10 % of the water demand to achieve "high resources autonomy" had a strong impact on the results in the Boom scenario. The stakeholder preferences (section 3.2) revealed that the objective of "high resource autonomy" is of negligible importance. Therefore, a variant of A6 which imports more water from external sources could be a viable option that performs better in the Boom scenario.

5.2 Preference elicitation

We find the presented two-step procedure consisting of an online survey and a face-to-face interview (section 2.2) very useful to elicit stakeholder preferences. The online survey helped to familiarize the ten stakeholders with the decision problem. Additionally, we obtained an importance ranking of the objectives that was used to focus interviews on the most relevant objectives and hence reduce complexity. The reasons for stakeholders to re-enter objectives in the interview that they had deemed "irrelevant" in the online survey are unclear and require more specific analyses. As reported in the literature, stakeholders often do not appropriately consider the attribute ranges, e.g. when stating weights (Morton & Fasolo, 2009; von Nietzsch & Weber, 1991 (in Eisenführ et al., 2010); Weber & Borcherding, 1993). This might have happened during the online survey, and recalling the ranges in the face-to-face situation might have led stakeholders to reconsider their initial judgment. Another reason could be that stakeholders felt uncomfortable with excluding potential taboo objectives such as "high social acceptance" in the face-to-face situation. A possible indication for this is that most interviewees gave very low weights to those re-included objectives. In MCDA-practice, online surveys to elicit preferences are rare, a major concern being the reliability of the responses and the introduction of biases if preferences are elicited unassisted from people that do not understand the implications (e.g. Marttunen & Hämäläinen, 2008; Mustajoki, Hämäläinen, & Marttunen, 2004). However, finding sound online survey procedures that allow eliciting detailed personal preferences required for MCDA, e.g. from a larger number of stakeholders, would be highly interesting. However, more (also experimental) research is needed (Lienert, Duygan, & Zheng, in preparation). We regard the suggested approach as one step in this direction.

The imprecision of stakeholder preferences was addressed by eliciting intervals (as suggested by e.g. Jessop, 2011; Mustajoki, et al., 2005) and a best guess. This was possible with minimal additional effort. Also, rough information about not-elicited value functions was obtained by asking just one preference difference question for each. In contrast to other shortened approaches reported in the literature (e.g. Schuwirth, et al., 2012), we did not ask for a quantitative specification of the equivalence point. This provides less detail, but can be elicited in considerably less time. Compared to detailed value function elicitation with the mid-value splitting method (taking about 15–45 min. for one value function), simplified elicitation was easy, worked instantly with all stakeholders, and was much faster (ca. 1–3 min. each). Results of the global sensitivity analysis (GSA) for SH2 revealed little influence of the uncertain value function shape on overall output uncertainty for three of five designs. Hence, for this case study, this approach provides a viable simplification for handling elicitation complexity and limited time without restricting the analysis to linear forms. In the case of the Boom scenario and when no acceptance thresholds apply, GSA gives clear indication on which of the uncertain value functions should be elicited in detail in order to reduce uncertainties.

During elicitation, strong preference thresholds regarding drinking water hygiene were identified, leading to the rejection of alternatives exceeding these thresholds (section 3.1, "Preference thresholds and individual adjustments", and 3.2). We did not find reports of similar experiences in practical MAUT applications. It remains open if stakeholders would have stated the same thresholds (in the case of drinking water hygiene even stricter than the predicted current situation), if more precise attribute predictions would have been already available at the time of the interviews. Precise knowledge of the attribute outcomes would have allowed one to choose less extreme attribute ranges, and contradictions of the stated acceptance thresholds for drinking water hygiene compared to the status quo could have been discussed with stakeholders. It was not possible to wait with detailed elicitation until the attribute predictions were completed, but for the above reasons we strongly encourage doing so in future studies. According to our experience with this case study, we also recommend to do consistency checks for acceptance thresholds. This seems advisable given their high impact on the results (demonstrated by the GSA).

We did not systematically assess the stakeholders' opinions about the elicitation methods themselves, but asked for feedback on the overall structured decision-making process this work was embedded into (Lienert, et al., 2014a; 2014b (Table 10)). The only issues with preference elicitation stated were: a) difficulty to understand some objectives and attributes ("good physico-chemical quality of drinking water"), b) missing objectives ("technology readiness", consideration of redundant water supplies e.g. in "high reliability of supply"), and c) doubts whether the full range of 0 up to 100 % end-user codetermination is desirable.

5.3 Uncertainty analysis

The uncertainty of preference parameters arising from imprecise statements, missing information about value and utility function curvatures, and the underlying aggregation functions was propagated to the outcomes of the alternatives (section 2.4). Also, the uncertainty of the attribute predictions was included. This goes much further than available uncertainty analyses (e.g. Butler, et al., 1997; Hyde, et al., 2004; Jessop, 2011; Jiménez, et al., 2006; Raju & Pillai, 1999), which mostly focus on the uncertainty of the weights and attributes only. It allowed important insights into the uncertainty of the resulting rankings, and how much these would deviate from an analysis under "usual simplification assumptions", namely linear single-attribute value functions, additive aggregation, sure weights (best-guess), and neutral risk attitude implied by neglecting uncertainty of attribute predictions (section 2.4). For some stakeholders, a strong divergence between the ranking of alternatives with uncertain preferences and the ranking for "usual simplification assumptions" was observed. For example, we would have recommended A6 or A1b given usual assumptions for SH1 in the Doom scenario, while the ranking under uncertain preferences resulted in better performance of A4 and A8b. The recommendations for potential compromise solutions for all stakeholders, however, were not affected. Despite this, the objective of the decision process might not always be to identify compromise solutions but rather to understand which alternatives are clearly best (or worst) for some stakeholders but not for others, and why. That can be important especially in deliberative processes with higher conflict potential than observed in this study. Although simplifications are clearly necessary for making elicitation and analysis feasible, the results caution us against oversimplification during preference modeling, because it may lead to wrong conclusions or recommendations. Nevertheless, the use of simplifying assumptions to explore (and not only to rank) alternatives may provide important insights. The process of helping stakeholders define their fundamental objectives and to use them to create and compare innovative solutions does not need the detailed consideration of uncertainties presented here (e.g. Gregory, et al., 2012). It might even be counterproductive for understanding due to being overwhelmed by uncertainties. On the other hand, the ignorance of uncertainties in current long-term planning is unlikely to be overcome by decision support approaches that do just the same, and the cost of being wrong due to its non-consideration can be high. This should be taken into account when MCDA evaluations are used to create legitimacy for particular decisions based on the ranking of alternatives.

The combination of uncertainty analysis and scenario planning was very beneficial for the comparison of alternatives in this case study. Whereas the ranking of alternatives in the Status quo, Quality of life, and Doom scenario were often similar, they diverged in the Boom scenario. These findings demonstrate that uncertain drivers of future change need to be considered in long-term water infrastructure planning (also advocated by e.g. Ferguson, et al., 2013; Milly, et al., 2008; Ruth, et al., 2007; Sharma, et al., 2010). The current narrow-minded extrapolation of the status quo under stationary assumptions (e.g. Ashley, et al., 2008; Dominguez, et al., 2009; Störmer, et al., 2009) should be overcome. The combination of MCDA and scenario planning provides a valuable framework for doing so, and furthermore allows including different stakes, which has often been overlooked in the past (e.g. Economides, 2012).

5.4 Global sensitivity analysis (GSA)

We also showed how GSA can be used to explore which of the uncertain parameters have the largest impact on the results (section 4). This information is helpful to better understand model behavior and to simplify elicitation for large objectives hierarchies. Textbooks require objectives hierarchies to be as concise as possible (e.g. Eisenführ, et al., 2010; Gregory, et al., 2012; Keeney & Raiffa, 1993), but there seems to be no consensus about how many objectives are still concise. Bond et al. (2008, 2010) found that unaided decision makers on average identify about seven relevant objectives, while they would identify twenty-two relevant objectives when picking from a master list. In practical MAVT / MAUT, it seems that only ten or less attributes are considered on average (Ananda & Herath, 2009; Hajkowicz & Collins, 2007; Mendoza & Martins, 2006). However, larger objectives hierarchies presumably often better reflect the complexity of real-world decision making (see e.g. Langhans, et al., 2013 for river rehabilitation). Therefore, we think there is a real need for viable solutions also in these more complex cases.

The results of the analysis indicate that our elicitation design (section 2.2) did focus on the most important parameters. Without any preference information available ("NoPref_SQ"), the results were highly sensitive to weights, both regarding their individual and interactive effects. This justifies the efforts spent on elicitation of the weights despite their large number (44 weights overall). Results from the other GSA layouts also imply that the remaining uncertainty from the weight intervals was negligible compared to the uncertainty about other parameters which were not elicited in detail and hence more uncertain. These parameters were chiefly the risk attitude (explaining up to 72 % of the output uncertainty in the case of stakeholder SH2; Status quo scenario), but also the aggregation form of the highest-level aggregation nodes, and the curvature of specific marginal value functions. We are not aware of any application regarding the elicitation of risk attitudes of multi-dimensional values. Usually, the curvatures of marginal (single-attribute) utility functions are elicited (e.g. Bleichrodt, et al., 2001; Eisenführ, et al., 2010; Keeney & Raiffa, 1993; Smidts, 1997). In our case, however, the precise elicitation of risk attitude might not even be necessary, despite its high influence. Since the value distributions are quite different for the better-performing alternatives, rough knowledge about whether the stakeholder is approximately risk-averse, risk-neutral, or risk-prone should be sufficient to order the alternatives. For example, if a stakeholder is risk-averse, alternatives that exceed acceptance thresholds (and therefore lead to zero value) can directly be excluded.

In all GSA layouts, more precise knowledge of a few parameters could strongly reduce ranking uncertainty. Efforts should thus be spent on eliciting these most important parameters. The challenge is, however, to determine the most influential parameters, as GSA computation is very costly. In this study, the uncertainty of the attribute predictions was only indirectly considered by calculating values and expected utilities for a fixed attribute sample. Therefore, we could not gain any insights regarding the importance of uncertainty in attribute predictions compared to the uncertainty of the preference parameters. In many cases, it will be important to know whether to spend efforts on obtaining better predictions or more detailed preference information. Additionally, improving the computational efficiency of GSA including many uncertain parameters e.g. as recommended by Saltelli et al. (2010) should be one objective of further studies, also to be able to obtain stable, quantitatively interpretable total order indices.

We would also like to raise attention to the high influence of interaction effects between parameters. Given the ubiquity of local (one at a time) sensitivity analyses in practical MAUT / MAVT and available software, we should be cautious when interpreting the influence of single parameters, because LSA is unable to capture interactive effects (e.g. discussed in Saltelli, et al., 2006).

6 Conclusions

We presented an approach to tackle uncertainty in a complex practical MAUT intervention and identified five major sources of uncertainty to be addressed during preference elicitation and modeling. These are the problem framing and structuring, attribute predictions, hierarchical aggregation function, marginal value or utility functions, and the weights. We explained how we dealt with these uncertainties in a complex case study on water supply infrastructure planning in Switzerland. A thorough uncertainty analysis was combined with a scenario planning approach regarding socio economic boundary conditions, to evaluate the performance of water supply infrastructure alternatives in light of uncertain preferences (and preference models), and uncertain attribute predictions for four future scenarios. Despite individually different preferences, we could identify potential compromise alternatives. The ranking of the alternatives changed most strongly under the highly dynamic Boom scenario, indicating that the consideration of changing boundary conditions (e.g. regarding population increase or decrease and the economic situation) is very important in long-term planning of water supply infrastructures.

Global sensitivity analysis (GSA) allowed us to assess the contribution of individual parameters and parameter groups on the uncertainty of the ranking of alternatives. In the presented example, the overall uncertainty in the ranking of alternatives can be largely reduced by additional elicitation of only a few parameters. An analysis assuming no preference information at all demonstrated in hindsight that our elicitation approach was able to address the most important uncertain preference parameters (in this case: the weights). It also showed that GSA can be helpful even prior to factual preference elicitation, to focus on reducing the uncertainty of those parameters which matter. To improve the presented elicitation approach, we suggest to split the interview into two parts. The first should be used to elicit interval weights and to check independence conditions. Also, rough information about the marginal value function forms can be obtained with the simplified procedure described herein. Based on these, a valid MAUT model is defined and an uncertainty analysis is done. If no clear ranking of alternatives can be derived from the results, GSA can be performed to determine which parameters should be elicited in-depth during a second, follow-up interview.

7 Acknowledgements

This research was funded within NRP 61 on Sustainable Water Management by the Swiss National Science Foundation (SNF), project number 406140_125901/1. We thank the stakeholders from the case study for their support and participation in workshops and interviews. We also thank numerous experts who supported performance estimation: Matthias Freiburghaus, Heiko Gebauer, Frederik Hammes, Juliane Hollender, Mark Honti, Thomas Hillenbrand, Daniel Hunkeler, Stefan Koetzsch, Eva Lieberherr, Ivana Logar, Christian Moeck, Kurt Nyffenegger, Andreas Peter, Mario Schirmer, Jürg Schmid, Stefan Schmid, Kurt Steiner, and Urs von Gunten. Special thanks go to our colleagues Christoph Egger, Katrin Wolter, Cristina Fritzsche, Max Maurer, Jun Zheng, and Anja Zorn for excellent collaboration and critical discussions. Fichtner IT Consulting is acknowledged for giving access to the FAST Asset Management software and Stefano Tarantola for his inputs on assessing the necessary size of the preference sample in GSA. Finally, we thank three anonymous reviewers for their helpful comments leading to the improvement of this manuscript.

Appendix

Table A1: Comparison of the performance of eleven alternatives (A1a to A9) in four future scenarios using usual simplifying assumptions and considering uncertain preferences. The individual ranking of alternatives given with usual simplifications (section 2.4), and summary statistics over all stakeholders using usual simplifications (mean of SH), and uncertain preferences (Mean rank of SH, median rank of SH) are shown. μ= mean, σ= standard deviation. Candidate best and worst compromises across stakeholders are bold and italicized. Individual mean / median ranks are shown in Table SI.11-12.

	USUAL SIMPLIFYING ASSUMPTIONS												UNCERTAIN PREFERENCES			
	Individual rankings										Mean of SH		Mean rank of SH		Median rank of SH	
	SH1	**SH2**	**SH3**	**SH4**	**SH5**	**SH6**	**SH7**	**SH8**	**SH9**	**SH10**	**μ**	**σ**	**μ**	**σ**	**μ**	**σ**
Boom																
A1a	5	3	5	2	3	3	3	3	3	5	3.5	1.0	3.9	0.61	3.9	0.88
A1b	6	2	2	1	2	2	2	2	1	6	2.6	1.7	2.6	1.21	2.7	1.34
A2	4	1	1	3	1	1	1	1	2	3	**1.8**	**1.1**	**2.1**	**0.99**	**2.2**	**1.32**
A3	10	9	10	9	10	10	9	10	9	9	9.5	0.5	9.6	0.61	9.6	0.70
A4	9	8	8	7	9	9	10	8	8	8	8.4	0.8	9.1	0.83	9.0	0.82
A5	11	10	11	10	11	11	11	11	10	11	**10.7**	**0.5**	**10.2**	**0.85**	**10.1**	**1.10**
A6	8	7	9	8	7	8	6	9	7	10	7.9	1.1	6.1	1.75	7.0	1.63
A7	7	6	7	6	8	7	8	7	6	4	6.6	1.1	5.8	1.30	6.4	1.07
A8a	2	5	4	5	5	5	5	5	5	2	4.3	1.2	4.2	1.59	4.0	1.56
A8b	1	4	3	4	4	4	4	4	4	1	3.3	1.2	3.4	1.53	3.1	1.60
A9	3	11	6	11	6	6	7	6	11	7	7.4	2.6	7.2	1.70	8.1	2.08
Doom																
A1a	3	4	3	3	3	6	3	4	2	8	3.9	1.7	4.7	1.08	4.4	1.43
A1b	2	2	1	2	2	2	2	3	1	7	2.4	1.6	3.0	1.40	2.7	1.70
A2	4	3	2	5	4	5	4	2	3	1	3.3	1.3	4.4	1.67	4.2	1.87
A3	10	9	10	9	10	10	9	10	9	11	9.7	0.6	**9.9**	**0.92**	**10.4**	**0.70**
A4	8	8	9	6	8	9	7	7	8	6	7.6	1.0	6.9	1.74	7.0	2.11
A5	11	10	11	10	11	11	11	11	10	10	**10.6**	**0.5**	9.6	1.13	9.9	0.74
A6	1	1	4	1	1	1	1	1	4	2	**1.7**	**1.2**	**2.3**	**1.26**	**2.1**	**1.60**
A7	9	5	8	8	9	7	10	9	7	9	8.1	1.4	6.7	2.09	6.9	2.33
A8a	6	6	6	4	6	3	5	5	5	4	5.0	1.0	4.7	1.20	4.5	1.35
A8b	5	7	5	7	5	4	6	6	6	3	5.4	1.2	5.4	1.77	5.1	1.66
A9	7	11	7	11	7	8	8	8	11	5	8.3	2.0	8.7	1.47	8.7	1.77
Quality of life																
A1a	1	4	2	2	3	6	3	4	2	7	3.4	1.8	4.5	0.93	4.3	0.95
A1b	2	2	1	3	1	2	2	3	1	8	2.5	2.0	3.5	1.75	2.9	2.18
A2	6	3	3	6	4	3	4	2	3	3	3.7	1.3	4.4	1.61	4.3	1.95
A3	10	9	10	9	10	10	10	10	9	11	9.8	0.6	**10.0**	**0.98**	**10.5**	**0.71**
A4	8	8	9	8	8	9	8	8	8	6	8.0	0.8	7.2	1.73	7.5	1.72
A5	11	10	11	10	11	11	11	11	10	10	**10.6**	**0.5**	9.5	1.15	9.7	0.82
A6	5	1	4	1	2	1	1	1	4	4	**2.4**	**1.6**	**2.7**	**1.50**	**2.4**	**1.80**
A7	9	5	8	7	9	7	9	9	7	9	7.9	1.3	6.4	1.96	6.8	1.81
A8a	4	6	6	4	6	5	5	6	5	1	4.8	1.5	4.5	1.45	4.5	1.43
A8b	3	7	5	5	5	4	6	5	6	2	4.8	1.4	4.8	1.95	4.5	1.58
A9	7	11	7	11	7	8	7	7	11	5	8.1	2.0	8.5	1.57	8.6	1.84
Status quo																
A1a	1	4	2	3	3	6	3	4	2	1	2.9	1.4	3.8	0.79	3.6	0.70
A1b	2	2	1	2	2	2	2	2	1	2	**1.8**	**0.4**	**2.4**	**0.97**	**2.1**	**1.20**
A2	6	3	3	4	4	3	4	3	3	3	3.6	0.9	4.6	1.58	4.7	1.77
A3	10	9	10	9	10	10	10	10	9	10	9.7	0.5	9.8	0.94	10.3	0.82
A4	7	8	9	6	8	9	7	7	8	8	7.7	0.9	7.0	1.69	7.2	1.99
A5	11	10	11	10	11	11	11	11	10	11	**10.7**	**0.5**	**9.6**	**1.20**	**9.8**	**0.92**
A6	3	1	4	1	1	1	1	1	4	4	2.1	1.4	2.7	1.63	2.4	1.90
A7	9	5	8	8	9	7	9	9	7	9	8.0	1.3	6.6	2.12	6.9	2.18
A8a	4	6	6	5	5	5	5	5	5	5	5.1	0.5	4.8	1.23	4.6	0.97
A8b	5	7	5	7	6	4	6	6	6	6	5.8	0.9	5.7	1.27	5.5	0.97
A9	8	11	7	11	7	8	8	8	11	7	8.6	1.6	8.8	1.36	9.0	1.49

References

Abdellaoui, M., Bleichrodt, H., & Paraschiv, C. (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science, 53*, 1659-1674.

Allais, M. (1953). Le comportement de l'homme rationel devant le risque: Critique des postulats et axiomes de l'ecole américaine. *Econometrica, 21*, 503-546.

Ananda, J., & Herath, G. (2009). A critical review of multi-criteria decision making methods with special reference to forest management and planning. *Ecological Economics, 68*, 2535-2548.

Ashley, R., Blackwood, D., Butler, D., Jowitt, P., Davies, J., Smith, H., Gilmour, D., & Oltean-Dumbrava, C. (2008). Making Asset Investment Decisions for Wastewater Systems That Include Sustainability. *Journal of Environmental Engineering-ASCE, 134*, 200-209.

Ayyub, B. M. (2001). *Elicitation of Expert Opinions for Uncertainty and Risks* (1st ed.): CRC Press LLC.

Belton, V., & Stewart, T. J. (2002). *Multiple criteria decision analysis: an integrated approach* Kluwer Academic Publishers.

Belton, V., & Stewart, T. J. (2010). Problem structuring for multiple criteria analysis. In M. Ehrgott (Ed.), *Trends in Multiple Criteria Decision Analysis* (pp. 209 - 239).

Ben-Akiva, M., & Lerman, S. (1989). *Discrete Choice Analysis*. USA: MIT Press.

Bleichrodt, H., Pinto, J. L., & Wakker, P. P. (2001). Making Descriptive Use of Prospect Theory to Improve the Prescriptive Use of Expected Utility. *Management Science, 47*, 1498-1514.

Bond, S. D., Carlson, K. A., & Keeney, R. L. (2008). Generating objectives: Can decision makers articulate what they want? *Management Science, 54*, 56-70.

Bond, S. D., Carlson, K. A., & Keeney, R. L. (2010). Improving the Generation of Decision Objectives. *Decision Analysis, 7*, 238-255.

Borcherding, K., Eppel, T., & von Winterfeldt, D. (1991). Comparison of Weighting Judgements in Multiattribute Utility Measurement. *Management Science, 37*, 1603-1619.

Butler, J., Jia, J., & Dyer, J. (1997). Simulation techniques for the sensitivity analysis of multi-criteria decision models. *European Journal of Operational Research, 103*, 531-546.

Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review, 18*, 139-165.

Cooke, R. M. (1991). *Experts in uncertainty: opinion and subjective probability in science*. New York: Oxford University Press.

Cox, J. C., Sadiraj, V., Vogt, B., & Dasgupta, U. (2012). Is there a plausible theory for decision under risk? A dual calibration critique. *Economic Theory*, 1-29.

Dominguez, D., Worch, H., Markard, J., Trujillo Alvarez, R., & Gujer, W. (2009). Closing the Capability Gap: Strategic Planning for The Infrastructure Sector. *California Management Review, 51*, 30-50.

Durbach, I. N., & Stewart, T. J. (2009). Using expected values to simplify decision making under uncertainty. *Omega, 37*, 312-330.

Durbach, I. N., & Stewart, T. J. (2011). An experimental study of the effect of uncertainty representation on decision making. *European Journal of Operational Research, 214*, 380-392.

Durbach, I. N., & Stewart, T. J. (2012a). A comparison of simplified value function approaches for treating uncertainty in multi-criteria decision analysis. *Omega, 40*, 456-464.

Durbach, I. N., & Stewart, T. J. (2012b). Modeling uncertainty in multi-criteria decision analysis. *European Journal of Operational Research, 223*, 1-14.

Dyer, J. S., & Sarin, R. K. (1979). Measurable Multiattribute Value Functions. *Operations Research, 27*, 810-822.

Dyer, J. S., & Sarin, R. K. (1982). Relative Risk Aversion. *Management Science, 28*, 875-886.

Economides, J. A. (2012). Public outreach: An engineer's perspective. *Journal - American Water Works Association, 104*, 74-76.

Eisenführ, F., Weber, M., & Langer, T. (2010). *Rational Decision Making* (1st ed.): Springer.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics, 75*, 643-669.

Ferguson, B. C., Frantzeskaki, N., & Brown, R. R. (2013). A strategic program for transitioning to a Water Sensitive City. *Landscape and Urban Planning, 117*, 32-45.

Franco, L. A., & Montibeller, G. (2011). Problem structuring for multicriteria decision analysis interventions. In J. J. E. Cochran (Ed.), *Encyclopedia of Operations Research and Management Science* (pp. 1-14). Oxford: John Wiley & Sons.

French, S. (1995). Uncertainty and Imprecision: Modelling and Analysis. *Journal of the Operational Research Society, 46*, 70-79.

French, S. (2003). Modelling, Making Inferences and Making Decision: The Roles of Sensitivity Analysis. *Sociedad de Estadística e Investigación Operativa (TOP), 11*, 229-251.
Gómez Delgado, M., & Bosque Sendra, J. (2004). Sensitivity Analysis in Multicriteria Spatial Decision-Making: A Review. *Human and Ecological Risk Assessment: An International Journal, 10*, 1173-1187.
Gómez Delgado, M., & Tarantola, S. (2006). Global sensitivity analysis, GIS and multi-criteria evaluation for a sustainable planning of a hazardous waste disposal site in Spain. *International Journal of Geographical Information Science, 20*, 449-466.
Goodwin, P., & Wright, G. (2001). Enhancing strategy evaluation in scenario planning: A role for decision analysis. *Journal of Management Studies, 38*, 1-16.
Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., & Ohlson, D. (2012). *Structured Decision Making- A practical guide to Environmental Management Choices* (1st ed.): Wiley-Blackwell.
Hajkowicz, S. (2008). A comparison of multiple criteria analysis techniques for water resource management. *European Journal of Operational Research,* , 255-265.
Hajkowicz, S., & Collins, K. (2007). A review of multiple criteria analysis for water resource planning and management. *Water Resources Management, 21*, 1553-1566.
Hershey, J. C., & Schoemaker, P. J. H. (1985). Probability versus certainty equivalence methods in utility measurement- are they equivalent? *Management Science, 31*, 1213-1231.
Hoeffler, S., & Ariely, D. (1999). Constructing Stable Preferences: A Look Into Dimensions of Experience and Their Impact on Preference Stability. *Journal of Consumer Psychology, 8*, 113-139.
Huang, I. B., Keisler, J., & Linkov, I. (2011). Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. *Science of the Total Environment, 409*, 3578-3594.
Hyde, K. M., Maier, H., & Colby, C. (2004). Reliability-based Approach to Multicriteria Decision Analysis for Water Resources. *Journal of Water Resources Planning and Management, 130*, 429-438.
Hyde, K. M., Maier, H. R., & Colby, C. B. (2005). A distance-based uncertainty analysis approach to multi-criteria decision analysis for water resource decision making. *Journal of Environmental Management, 77*, 278-290.
Jacobi, S. K., & Hobbs, B. F. (2007). Quantifying and Mitigating the Splitting Bias and Other Value Tree-induced Weighting Biases. *Decision Analysis, 4*, 194-210.
Jessop, A. (2011). Using imprecise estimates for weights. *Journal of the Operational Research Society, 62*, 1048-1055.
Jiménez, A., Ríos-Insua, S., & Mateos, A. (2006). A generic multi-attribute analysis system. *Computers & Operations Research, 33*, 1081-1101.
Joubert, A., Stewart, T. J., & Eberhard, R. (2003). Evaluation of Water Supply Augmentation and Water Demand Management Options for the City of Cape Town. *Journal of Multi-Criteria Decision Analysis*, 17-25.
Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica, 47*, 263-291.
Kangas, A. S., & Kangas, J. (2004). Probability, possibility and evidence: approaches to consider risk and uncertainty in forestry decision analysis. *Forest Policy and Economics,* , 169-188.
Karvetski, C. W., Lambert, J. H., & Linkov, I. (2009a). Emergent conditions and multiple criteria analysis in infrastructure prioritization for developing countries. *Journal of Multi-Criteria Decision Analysis, 16*, 125-137.
Karvetski, C. W., Lambert, J. H., & Linkov, I. (2009b). Scenario and multiple criteria decision analysis for energy and environmental security of military and industrial installations. *Integrated Environmental Assessment and Management, 16*, 125-137.
Keeney, R. L. (1982). Decision Analysis: An Overview. *Operations Research, 30*, 803-838.
Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives. Preferences and Value tradeoffs* (2 ed.). Cambridge: Cambridge University Press.
Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika, 30*, 81-93.
Kilgour, D. M., & Eden, C. (2010). Handbook of Group Decision and Negotiation. In. Dordrecht, Heidelberg, London, New York: Springer.
Kynn, M. (2008). The 'heuristics and biases' bias in expert elicitation. *Journal of the Royal Statistical Society: Series A (Statistics in Society), 171*, 239-264.
Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA - Stochastic multiobjective acceptability analysis. *European Journal of Operational Research, 106*, 137-143.
Lahdelma, R., & Salminen, P. (2012). The shape of the utility or value function in stochastic multicriteria acceptability analysis. *OR Spectrum, 34*, 785-802.
Langhans, S. D., Lienert, J., Schuwirth, N., & Reichert, P. (2013). How to make river assessments comparable: A demonstration for hydromorphology. *Ecological Indicators,* , 264-275.

Langhans, S. D., Reichert, P., & Schuwirth, N. (2014). The method matters: indicator aggregation in ecological assessment. *Ecological Indicators, 45*, 494-507.

Lieberherr, E., Klinke, A., & Finger, M. (2012). Towards legitimate water governance? The partially privatized Berlin waterworks. *Public Management Review, 14*, 923-946.

Lienert, J., Duygan, M., & Zheng, J. (in preparation). Preference stability over time using two methods to elicit MCDA-weights for wastewater infrastructure planning (working title).

Lienert, J., Schnetzer, F., & Ingold, K. (2013). Stakeholder analysis combined with social network analysis provides fine-grained insights into water infrastructure planning processes. *Journal of Environmental Management, 125*, 134-148.

Lienert, J., Scholten, L., Egger, C., & Maurer, M. (2014a). Structured decision making for sustainable water infrastructure planning and four future scenarios. *EURO Journal on Decision Processes*, http://dx.doi.org/10.1007/s40070-40014-40030-40070.

Lienert, J., Scholten, L., Egger, C., & Maurer, M. (2014b). Additional information for Structured decision making for sustainable water infrastructure planning and four future scenarios. In *Working Papers in Environmental Social Sciences 2014-3, Department of Environmental Social Sciences, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland.*http://www.eawag.ch/forschung/ess/workingpapers/. *Accessed 06 December 2013.*

Martin, W. E., Bender, H. W., & Shields, D. J. (2000). Stakeholder objectives for public lands: Rankings of forest management alternatives. *Journal of Environmental Management,* , 21-32.

Marttunen, M., & Hämäläinen, R. P. (2008). The Decision Analysis Interview Approach in the Collaborative Management of a Large Regulated Water Course. *Environmental Management,* , 1026-1042.

Mendoza, G. A., & Martins, H. (2006). Multi-criteria decision analysis in natural resource management: A critical review of methods and new modelling paradigms. *Forest Ecology and Management*, 1-22.

Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., & Stouffer, R. J. (2008). Climate change – Stationarity is dead: Whither water management? *Science, 319*, 573-574.

Montibeller, G., Gummer, H., & Tumidei, D. (2006). Combining scenario planning and multi-criteria decision analysis in practice. *Journal of Multi-Criteria Decision Analysis, 14*, 5-20.

Morton, A., & Fasolo, B. (2009). Behavioural decision theory for multi-criteria decision analysis: a guided tour. *Journal of the Operational Research Society, 60*, 268-275.

Mullen, K. M., Ardia, D., Gil, D. L., Windover, D., & Cline, J. (2011). DEoptim: An R package for Global Optimization by Differential Evolution. *Journal of Statistical Software, 40*.

Mustajoki, J. (2011). Comparison of weight distributions in multi-attribute value tree theory. In V. Kostaglou, G. Arabatzis & L. Karamitopoulos (Eds.), *International Symposium & 10th Balkan Conference on Operational Research (BALCOR 2011)* (Vol. 1, pp. 300-307). Thessaloniki, Greece.

Mustajoki, J. (2012). Effects of imprecise weighting in hierarchical preference programming. *European Journal of Operational Research, 218*, 193-201.

Mustajoki, J., Hämälainen, R. P., & Lindstedt, M. R. K. (2006). Using intervals for global sensitivity and worst-case analyses in multiattribute value trees. *European Journal of Operational Research, 174*, 278-292.

Mustajoki, J., Hämäläinen, R. P., & Marttunen, M. (2004). Participatory multicriteria decision analysis with Web-HIPRE: a case of lake regulation policy. *Environmental Modelling & Software, 19*, 537-547.

Mustajoki, J., Hämäläinen, R. P., & Salo, A. (2005). Decision support by interval SMART/SWING - Incorporating imprecision in the SMART and SWING methods. *Decision Sciences, 36*, 317-339.

Nash, J. C., & Varadhan, R. (2011). Unifying Optimization Algorithms to Aid Software System Users: optimx for R. *Journal of Statistical Software, 43*, 1-14.

O'Hagan, A., Buck, C. E., Daneshkhah, A., Eiser, J. R., Garthwaite, P. H., Jenkinson, D. J., Oakley, J. E., & Rakow, T. (2006). *Uncertain Judgements: Eliciting Experts' Probabilities*: John Wiley & Sons.

Phillips, L. D., & Bana e Costa, C. A. (2007). Transparent Prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Annals of Operations Research, 154*, 51-68.

Pöyhönen, M., Vrolijk, H., & Hämäläinen, R. P. (2001). Behavioral and procedural consequences of structural variation in value trees. *European Journal of Operational Research, 134*, 216-227.

Pujol, G., Iooss, B., & Janon, A. (2012). Sensitivity: Sensitivity Analysis. In (pp. R package version 1.6-1).

Qualtrics. (2012). Qualtrics Research Suite. In (2.227s ed.). Pravo, Utah, USA.

R Development Core Team. (2011). R: A Language and Environment for Statistical Computing. In. Vienna, Austria: R Foundation for Statistical Computing.

Rabin, M. (2000). Risk Aversion and Expected-Utility Theory: A Calibration Theorem. *Econometrica, 68*, 1281-1292.

Raju, K. S., & Pillai, C. R. S. (1999). Multicriterion decision making in performance evaluation of an irrigation system. *European Journal of Operational Research,* , 479-488.

Raju, K. S., & Vasan, A. (2007). Multi attribute utility theory for irrigation system evaluation. *Water Resources Management, 21*, 717-728.

Refsgaard, J. C., van der Sluijs, J. P., Højberg, A. L., & Vanrolleghem, P. A. (2007). Uncertainty in the environmental modelling process - A framework and guidance. *Environmental Modelling & Software, 22*, 1543-1556.

Reichert, P., Schuwirth, N., & Langhans, S. D. (2013). Constructing, evaluating, and visualizing value and utility functions for decision support. *Environmental Modelling & Software, 46*, 283-291.

Ruth, M., Bernier, C., Jollands, N., & Golubiewski, N. (2007). Adaptation of urban water supply infrastructure to impacts from climate and economic changes: The case of Hamilton, New Zealand. *Water Resources Management,* , 1031-1045.

Saltelli, A. (2008). *Global sensitivity analysis: the primer*. Chichester: John Wiley.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications, 181*, 259-270.

Saltelli, A., Ratto, M., Tarantola, S., & Campolongo, F. (2006). Sensitivity analysis practices: Strategies for model-based inference. *Reliability Engineering and System Safety, 91*, 1109-1125.

Saltelli, A., Tarantola, S., Campolongo, F., & Ratto, M. (2004). *Sensitivity Analysis In Practice: A Guide to Assessing Scientific Models*: John Wiley & Sons Ltd.

Saltelli, A., Tarantola, S., & Chan, K. (1999a). A role for sensitivity analysis in presenting the results from MCDA studies to decision makers. *Journal of Multi-Criteria Decision Analysis, 8*, 139-145.

Saltelli, A., Tarantola, S., & Chan, K. P. S. (1999b). A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics, 41*, 39-56.

Scheidegger, A. (2012). adaptMCMC: Implementation of a generic adaptive Monte Carlo Markov Chain sampler. In (v1.0.3 ed.): CRAN R project.

Schmidt, U., Starmer, C., & Sugden, R. (2008). Third-generation prospect theory. *Journal of Risk and Uncertainty, 36*, 203-223.

Schuwirth, N., Reichert, P., & Lienert, J. (2012). Methodological aspects of multi-criteria decision analysis for policy support: A case study on pharmaceutical removal from hospital wastewater. *European Journal of Operational Research, 220*, 472-483.

Sharma, A., Burn, S., Gardner, T., & Gregory, A. (2010). Role of decentralised systems in the transition of urban water systems. *Water Science and Technology: Water Supply, 10*, 577-583.

Smidts, A. (1997). The relationship between risk attitude and strength of preference: A test of intrinsic risk attitude. *Management Science, 43*, 357-370.

Stewart, T. J. (1995). Simplified approaches for multicriteria decision making under uncertainty. *Journal of Multi-Criteria Decision Analysis, 4*, 246-258.

Stewart, T. J. (2005). Dealing with uncertainties in MCDA (Multi-criteria decision analysis). In J. Figueira, S. Greco & M. Ehrgott (Eds.), *Multiple Criteria Decision Analysis - State of the Art Annotated Surveys* (Vol. 78, pp. 445 - 470). New York: Springer.

Stewart, T. J., French, S., & Rios, J. (2013). Integrating multicriteria decision analysis and scenario planning—Review and extension. *Omega, 41*, 679-688.

Störmer, E., Truffer, B., Dominguez, D., Gujer, W., Herlyn, A., Hiessl, H., Kastenholz, H., Klinke, A., Markard, J., Maurer, M., & Ruef, A. (2009). The exploratory analysis of trade-offs in strategic planning: Lessons from Regional Infrastructure Foresight. *Technological Forecasting and Social Change,* , 1150-1162.

Torrance, G. W., Feeney, D. H., Furlong, W. J., Barr, R. D., Zhang, Y., & Wang, Q. (1996). Multiattribute Utility Function for a Comprehensive Health Status Classification System: Health Utilities Index Mark 2. *Medical Care, 34*, 702-722.

von Neumann, J., & Morgenstern, O. (1947). *Theory of Games and Economic Behavior* (2 ed.): Princeton University Press.

von Nietzsch, R., & Weber, M. (1991). Bandbreiten-Effekt bei der Bestimmung von Zielgewichten. *Zeitschrift für betriebswirtschaftliche Forschung, 43*, 971-986.

Wakker, P., & Deneffe, D. (1996). Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. *Management Science, 42*, 1131-1150.

Walker, W. E., Harremoës, P., Rotmans, J., van der Sluijs, J. P., van Asselt, M. B. A., Janssen, P., & von Krauss, M. P. K. (2003). Defining Uncertainty: A Conceptual Basis for Uncertainty Management in Model-Based Decision Support. *Integrated Assessment, 4*, 5 - 17.

Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., & Deb, K. (2008). Multiple criteria decision making, multiattribute utility theory: Recent accomplishments and what lies ahead. *Management Science, 54*, 1336-1349.

Weber, M. (1987). Decision making with incomplete information. *European Journal of Operational Research*, 44-57.

Weber, M., & Borcherding, K. (1993). Behavioral influences on weight judgments in multiattribute decision-making. *European Journal of Operational Research, 67*, 1-12.

Wickham, H. (2007). Reshaping data with the reshape package. *Journal of Statistical Software, 21*.

Wickham, H. (2009). ggplot2: elegant graphics for data analysis. In: Springer New York.

Wong, T. H. F., & Brown, R. R. (2009). The water sensitive city: Principles for practice. *Water Science and Technology, 60*, 673-682.

Zhou, P., & Ang, B. W. (2009). Comparing MCDA Aggregation Methods in Constructing Composite Indicators Using the Shannon-Spearman Measure. *Social Indicators Research, 94*, 83-96.

Tackling uncertainty in multi-criteria decision analysis – An application to water supply infrastructure planning

List of symbols and abbreviations

Symbol/ abbreviation	Explanation
a	alternative
AT	acceptance threshold
c_j	marginal value function curvature over attribute j, j=1…30
EU(a)	expected utility of an alternative a
GSA	global sensitivity analysis
LSA	local sensitivity analysis
n	attribute sample size
$p_a(x)$	probability density of the attributes x for alternative a
r	curvature of the utility function
s	preference parameter sample size
S_{Tz}	total order sensitivity coefficient (due to interactions)
S_z	first order sensitivity coefficient (due to individual main effect)
U(v)	hierarchical (multi-attribute)
V(v)	hierarchical (multi-attribute)
v_i	marginal value of alternative regarding objective i
w_i	weights, i=1…44
x_j	attribute level x of attribute j
α_k	aggregation mixture parameter, k = 1…15; $\alpha_k \in [0,1]$
θ_z	vector of z parameters θ_z, $z = 1 \dots 90$

SI1 Stakeholder identification

More details about the underlying stakeholder (SH) and social network analysis (SNA) are given in (Lienert et al., 2013). The meaning of the selection criteria is as follows:

- Influence on infrastructure planning: Interviewees rated the strength of the influence of a SH on water infrastructure planning on a 0 – 10 scale (0: no influence; 10: cannot make infrastructure decisions without). The mean of all interviews was used.

- Affectedness by infrastructure planning: Identical to above, but assessing how strongly a stakeholder is affected once a decision is made; from 0 (not at all) to 10 (very strongly).

- Maximum number of times mentioned in interviews to influence or be affected by infrastructure planning (e.g. if 27 = SH was mentioned at least once in each of 27 interviews).

- Ability to overcome barriers in infrastructure planning: Number of times a SH was mentioned.

- Providers of resources for infrastructure planning: Number of times a SH was mentioned.

- Degree of centrality: This term from social network analysis describes the structural importance of a SH within the SH network. The degree centrality takes the ties an actor directly shares with the other actors into account and looks at the local structure she or he is embedded in. SHs with high degree centrality have better and direct access to information and have the potential to frame the planning process considerably.

- Betweenness centrality: This term from social network analysis assesses the power and importance of a SH derived from how often he or she is on the path between two SHs which are not linked to each other. A SH with high betweenness centrality can act as a 'gatekeeper' or mediator and are important for maintaining the network.

- Location within stakeholder network: This term from social network analysis describes how central or peripheral a SH is to the social network. If in the core (1), then SH is central to the network (= important, primary role in infrastructure planning), otherwise at the periphery (=0, secondary role).

Table SI1.1: Importance of selected stakeholders for infrastructure planning (ISP) based on 27 face-to-face interviews (Lienert et al., 2013). SH = stakeholder; ISP = infrastructure planning; SNA = social network analysis

SH	Stakeholder	Influence on ISP	Affectedness by ISP	Frequency mentioned in interviews	Ability to overcome barriers in ISP	Providers of resources for ISP	Degree of centrality (SNA)	Betweenness centrality (SNA)	Location within stakeholder network (SNA)
1	**Municipal underground engineer**	7.4	5.6	15	5	27	0.275	0.430	1
2	**Operating staff**	6.4	6.6	7	0	6	0.175	0.285	1
3	**Local water supply cooperative**	5.7	6.8	9	2	20	0.275	0.404	1
4	**Municipal administration***	6.0	7.1	6	0	24	0.300	0.435	1
4	**Municipal engineering and finance***	7.6	6.3	9			0.325	0.523	1
5	**Engineering consultant**	6.7	6.4	12	0	3	0.225	0.369	1
6	**Regional water supply cooperative**	5.4	5.4	4	0	5	0.175	0.335	1
7	**Cantonal environmental protection agency**	5.2	4.8	13	7	10	0.200	0.300	1
8	**Cantonal (water) quality laboratory**	6.1	4.6	6	0	1	0.175	0.281	1
9	**National association of gas and water industry**	3.2	1.7	7	5	6	0.05	0.103	0
10	**National environmental protection agency**	1.1	0.5	5	1	2	0.100	0.091	0

* The positions of 'municipal administration' and 'municipal engineering and finance' are shared by the same individual.

SI2 Decision attributes

Description of attributes and quantification

Table SI2.1: Overview of the ranges, description, and assessment of the attributes. The distributions and corresponding parameters used are shown in Table SI3.2. For more details also see Lienert et al. (submitted).

Short name	Attribute [unit]	Range ($^+$: best level)	Description	Assessment
rehab	Realization of the rehabilitation demand [%]	0-100$^+$	In the short term, purely repair-based rehabilitation strategies are cheaper than renewal or replacement strategies. The consequence is a water infrastructure which not only has a higher average age, but which is also more prone to failure. Undetected leakage leads to high increased water losses.	Calculated. Rehabilitation of the centralized pipe water system is modeled in detail following the approach described in (Scholten et al., 2014). The therein specified prior distribution is used to predict failures for the case study networks as a whole, but without Bayesian inference of failure parameters (because there are no failure records from three of the five case study water networks and because of the little difference between the prior and posterior distribution shown in (Scholten et al., 2014) for water supplier D). The replacement of treatment, pumping, and storage facilities of the centralized and decentralized treatment system are not considered given their much shorter lifetimes and higher immediacy. Partial replacements are often performed during usual maintenance. For these assets, a 100% realization of the rehabilitation demand within one generation is assumed.
adapt	Flexibility of technical extension or deconstruction of infrastructure [%]	0-100$^+$	A measure indicating how easy it is to technically extend or deconstruct the infrastructure. This depends on organizational structure, construction and operation of infrastructure, and drinking water system technology.	Expert assessment. At first, all alternatives were judged individually by four participating engineers. Their judgment was incurred concerning how easy it would be to technically extend or to deconstruct the respective infrastructure. Thereto each participant received a form with a description of the relevant aspects characterizing the alternatives, namely: organizational structure, construction and operation of water infrastructure, wastewater system technology, and drinking water system technology. The participant assigned one out of the five categories "very low (0-20%)", "low (20- 40%)", "medium (40- 60%)", "high (60- 80%)", "very high (80-100%) system flexibility" to each alternative. Then, the mean of the participants' judgments and the standard deviation were calculated (using the mid-points of the categories' intervals, i.e., 10, 30, 50, 70, and 90%). Those alternatives with more than 10% deviation were subsequently discussed. The group members with the highest divergence explained the argumentation for their judgments. After this was done, a final score was assigned to each alternative by the overall group. Larger interval ranges depict higher uncertainty or higher variance between the group member's judgments. These results were sent to two external experts (Fraunhofer Institute for Systems and Innovation Research ISI, Karlsruhe, Germany; Institute for social-ecological research ISOE, Frankfurt, Germany) for validation.
gwhh	% Utilization of groundwater	$^+$0-180	Raw water can be abstracted from springs and groundwater wells in the	Calculated as groundwater abstraction/groundwater recharge. Groundwater recharge was estimated using the Hydrus1D model for simplified soil profiles, representing

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
	recharge [%]		region, or imported from other sources (e.g. lake water from regional water supplier). The environmental sustainability of the groundwater balance is linked to the proportion of abstracted groundwater in comparison to the amount of natural groundwater recharge (e.g. from rain).	the characteristics of predominant soils in the case study region. Climate data (MeteoSwiss, 2011) and delta change scenarios for ten different regional climate models were used (Bosshard et al., 2011; CH2011, 2011). Based on these, rain series were generated in a collaboration project (iWaQa, 2013, personal communication) using a weather generator (Kilsby et al., 2007) following the description of (Fatichi et al., 2011). The minimum and maximum resulting range for groundwater recharge per m^2 was used. The political area of the case study is used as a reference, i.e. groundwater abstraction and recharge are calculated as per m^2 of political land area. The amount of groundwater abstraction depends on the scenario and alternative.
econs	Net energy consumption for water treatment and transport [kWh/m^3]	+0-2	Energy consumption depends on how the water is treated and transported to the end users (i.e. the particular treatment installations, the amount of pumping requires or the km distance covered by lorry transport).	Calculated. The best case (low energy consumption) is assumed to be zero, because of little / no treatment of water and wastewater, and the use of gravity for transport. The worst case (maximum energy consumption) was calculated assuming very energy-intensive water treatment, and water withdrawal and transport over long distances requiring pumps and tank wagons. To transport bottled water, mineral oil equivalents were converted to energy. For wastewater, we assumed the energy consumption of high tech decentralized treatment units, and added the energy consumption for the removal of micropollutants and the treatment of urine (and a safety factor). Energy demand for water treatment and distribution is calculated based on assumptions from (Vince et al., 2008) for different centralized treatment and distribution systems. Energy demand for advanced oxidation processes origins from (Katsoyiannis et al., 2011). Energy for household pumping and treatment is calculated according to producer specifications of the selected decentralized installations. The energy demand for water lorries is taken from (TREMOD, 2010). Bottled water is presumably bought together with other goods and thus its impact regarding energy (fuel) consumption was neglected.
vol_dw, vol_hw, vol_ffw	Days per year with water quantity limitations [d/a] FFW: Available water for firefighting in new housing areas [L/min]	+0-365 500-3600+	Quantity limitations as regards the water source are not expected because of different water sources available in the region (besides local springs and groundwater sources, vast reserves of lake water exist). Hence, water quantity limitations here refer to those induced by a mismatch of the technically dimensioned supply capacity and the demand.	Calculated. Whether a system is prone to water quantity limitations or not depends on the dimensioning size of the system and the expected demand. Centralized pipe systems were dimensioned on peak demands and are thus less prone to quantity limitations than decentralized tank options dimensioned on satisficing average daily demands. Following explanations of one of the local engineering consultants, the peak hourly demand currently used for dimensioning amounts to 450 L/(inhabitant*d) which is considerably less than the amounts used in the past (around 550-800 L/(inh.*d), population-weighted mean ca. 640 L/(inh.*d)), but sufficient to cover past residential peak demands in the case study water networks. Only in the network of one water supplier, the peak measured demand over the last decade is 471.4 L/(Ed). Except of this single event, on 99.7% of days between 2007-2010, the water demand amounted to less than 390.3 L/(inh.*d). Hence it is assumed, that the centralized pipe network is not likely to expect water quantity restrictions if dimensioned to that peak demand (450 L/(inh.*d), peak hour demand = 10% of peak days). If the decentralized systems are delivery on demand systems

Short name	Attribute [unit]	Range ($^{+}$: best level)	Description	Assessment
				(or buying water in the supermarket), it is also assumed that quantity limitations are unlikely. In the case of alternative A4, in the Boom scenario, water is refilled in regular, weekly intervals. Using the rain time series generated for the predictions of groundwater recharge (see gwhh) and assuming a completely filled rainwater tank at the beginning, the number of days with quantity restrictions are counted.
reliab_dw, reliab_hw, reliab_ffw	System reliability (in interviews termed "criticality") [-]	$^{+}$0-0.25	Assessment was done using the term "criticality", not reliability as is now used for correctness. The reliability is a dimensionless index which describes how many interruptions of service of what strength are expected. Assets of higher criticality (e.g. large pipes) receive a higher criticality weight, than assets with lower criticality (e.g. small pipes).	Calculated. The estimates of system reliability are based on the probability of failure, which is modeled in detail for the centralized pipe system and the criticality of different assets. In decentralized systems, a discrete scale is used. As orientation, the classification of failure rates in decentralized wastewater systems as reported in (Jones et al., 2004) is used. It classifies the annual probability of failure as associated to a qualitative judgment from very high (failure rate (FR): >1-1) over high (FR: 0.5-0.33), moderate (FR: 0.25-0.1), and rare (0.05-0.03) to extremely rare (FR 0.02).
aes_dw, aes_hw	Days per year with esthetic impairment such as taste, smell, etc.[d/a]	$^{+}$0-365	Water quality can be impaired due to different reasons, mainly smell, taste, discoloring, and turbidity. The aesthetics depend on the characteristics of the raw water and the technical installations (quality and type of water purification, dimensioning regional stagnations, operation, and maintenance).	Expert assessment. An expert from the Zurich cantonal laboratory provided the estimates. Thereto, two meetings were convened, before the expert assessed the alternatives. In the first meeting, characteristics of the case study area, the alternatives, and the future scenarios were presented and discussed. Factors that influence the attribute were discussed. The expert defined which additional information he needed to provide estimates for the attribute levels. In the second meeting, the requested additional information and detailed characteristics of the alternatives were presented and discussed.
faecal_dw, faecal_hw	Days per year with hygienic concerns (hygiene indicators) [d/a]	$^{+}$0-365	By law, drinking water must be free of organisms of hygienic concern, but their occurrence is not impossible. Indicator organisms ("fecal indicators") are used to test water. Reasons for the occurrence of fecal indicators can be inadequate purification, long stagnation in the network, inappropriate cleaning of the system, or pollution caused by misconnected pipes.	Expert assessment. Expert and assessment as for aes_dw, aes_hw.
cells_dw, cells_hw	Changes in total cell count as indicator of bacterial regrowth [log units]	$^{+}$0-2	Cell counts are indicators for the amount of microorganisms in water and serve to monitor bacterial regrowth in water supply systems. Distinction between active and inactive organisms is currently not possible. Every system has an equilibrium concentration of cells. Changes in cell counts indicate changes	Expert assessment. Expert and assessment as for aes_dw, aes_hw, but with an additional estimate of an expert at Eawag (specialist in flow cytometric cell counts). The estimate of both experts were combined, i.e. the overall average, maximum, and minimum values were used.

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
			in the microbial community and hence regrowth, which is usually of higher interest than absolute cell counts.	
no3_dw, no3_hw	Inorganic substances (indicator: nitrate concentration) [mg/L]	+0-20	Although nitrate itself is not toxic to humans unless occurring in much higher concentrations, European drinking water regulations decided to keep the levels below 50 mg/L (40 mg/L in Switzerland) for precautionary health reasons as nitrate can be used as general indicator parameter for other possibly toxic or carcinogenic nitrogen compounds (e.g. nitrite, nitrosamines). The Swiss water protection directive (GSchV, 2011) is limiting it to less than 25 mg/L mostly out of ecological considerations.	Attribute ranges. Time did not suffice to estimate this attribute in detail. Hence, the minimum and maximum attribute ranges are used. These stem from the measured concentrations in the different raw waters in the case study region (AWEL, 2013) and lake water at Stäfa (Stadt Zürich, 2012), and the minimum and maximum mixing ratios of these. It is assumed that some treatment can be found which might lead to a complete removal of nitrate.
pest_dw, pest_hw	Pesticides (sum of pesticide concentration) [μg/L]	+0-0.02	The sum of pesticides can be used as indicator parameter for agricultural and urban activities in the raw water catchment area. For precautionary environmental and health reasons, drinking water regulations in Switzerland demand the sum of pesticides to be below 0.5 μg/L and less than 0.1 μg/L for individual substances (FIV, 2009).	Attribute ranges. Time did not suffice to estimate this attribute in detail. Hence, the minimum and maximum attribute ranges are used. These stem from the measured concentrations in the different raw waters in the case study region (AWEL, 2013) and lake water at Stäfa (Stadt Zürich, 2012) and the minimum and maximum mixing ratios of these. It is assumed that some treatment can be found which might lead to a complete removal of pesticides.
bta_dw, btw_hw	Micropollutants (indicator: benzotriazole) [ng/L]	+0-150	Benzotriazole is a micropollutant used in coolants, for corrosion protection of surfaces, or de-icing purposes. Due to its high water solubility, limited sorption tendency, and low degradability, it is one of the most ubiquitous micropollutants observed in the Swiss environment. To avoid adverse health effects to the natural ecosystems and humans, the maximum recommended discharge concentrations for wastewater are 120 μg/L for single-discharge events and 30 μg/L for chronic discharges. Appropriate thresholds for toxicological concern in drinking water	Attribute ranges. Time did not suffice to estimate this attribute in detail. Hence, the minimum and maximum attribute ranges are used. These stem from the measured concentrations in the different raw waters and the minimum and maximum mixing ratios of these. It is assumed that some treatment can be found which might lead to a complete removal of benzotriazole.

Short name	Attribute [unit]	Range ($^+$: best level)	Description	Assessment
			are yet under discussion.	
efqm	Score of the EFQM excellence model (European Foundation for Quality Management) [%]	20-95$^+$	The EFQM Excellence Model is used to assess the quality of operations and management. Assessment is based on the organizational form and the spatial extent of the alternatives.	Expert assessment. For details concerning the model see (EFQM, 2012). An expert from Eawag provided the estimates. The same procedure as in the case of aes_dw, aes_hw, cells_dw, cells_hw was followed. Through nine criteria, the EFQM Excellence Model helps companies understand and analyze the cause and effect relationships between what the organization does and the results it achieves. Five of these criteria are 'Enablers' and four are 'Results'. The 'Enabler' criteria cover what an organization does and how it does it. The 'Results' criteria cover what an organization achieves (EFQM 2012). Each alternative is assessed separately, assigning up to 100 points each and then normalized to a range of 0-100**%**. The “results” criteria were discarded as the expert judged a fictitious judgment of future results based on organization from and spatial extent pointless.
voice	Degree (percent) of codetermination [%]	0-100$^+$	Describes how much end users have a say in water infrastructure decisions. Relevant influences are the organizational structure, geographic extent, and financial strategy.	Expert assessment. Two experts from Eawag provided the estimates. After information and discussion about the alternatives and future scenarios, all alternatives were judged individually by the expert. They assigned one of five categories “very low (0- 20%)”, “low (20- 40%)”, “medium (40- 60%)”, “high (60- 80%)”, “very high (80- 100%) system codetermination” to each alternative. The estimates of both experts were integrated to get an overall minimum, maximum, and average value.
auton	% of water coming from the Mönchaltorfer Aa region [%]	0-100$^+$	The more water originates from the region, the more autonomy decision makers have about its use. It is described by how much of the water used in the case study area stems from tertiary parties outside the case study.	Calculated. The percentage of water abstracted from sources and wells in the case study region depends on the alternative and the water demand. The water demand covers household, industry, and business demand as well as water losses. It is calculated depending on the future scenario and the alternative. Water which is imported from the regional water supply cooperative (surface water from lake Zurich) is considered ‘external’ and reducing resources autonomy of the case study area.
time	Necessary time investment for operation and maintenance by user [h/(inh.*a)]	$^+$0-10	This attribute estimates the time each citizen has to invest per year to operate and maintain decentralized water supply installations. This can involve e.g. the cleaning or exchange of filters, or the maintenance of tanks. Also telephone calls to ask for help by a specialist require time.	Calculated. Only applies to decentralized installations in private households which the end user takes care of. Necessary operation and maintenance times depend on the water supply facilities as specified by the alternative and following dimensioning for different building units. Time demands are specified by installation and building unit, added up and then divided by the number of inhabitants sharing a unit. Building units are areas of approximately similar housing and density. The existing building areas in the case study were summarized into 10 building units, 5 for the Status quo/Doom scenario, 3 for the Boom scenario, 2 for the Quality of lifeQuality of life scenario. A weighted mean over all building units is calculated for estimation.
area	Additional area demand on private property per end user [m^2/inh.]	$^+$0-10	Decentralized water supply systems such as decentralized tanks or point-of-entry or point-of-use treatment in households require additional space on private	Calculated. Only applies to decentralized installations in private households with additional space needs. The different installations are dimensioned for predefined building units (see explanation under “time”) and then the area demand for each building unit can be calculated. The area per building unit is divided by the number

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
			ground.	of inhabitants in the building unit and a weighted mean calculated over all building units in the case study area.
collab	Number of infrastructure sectors that collaborate in planning and construction [-]	1-6+	This attribute judges for each of the decision alternatives in SWIP, how many of six sectors that use the underground collaborate. As an example, if the drainage company is renewing its sewers in a specific section, the gas and water infrastructure rehabilitation could also be carried out together. Otherwise it could happen that right after the construction works of one sector, another sector starts its amelioration works, hereby reopening practically the same "hole".	Direct consequence of the alternative definition. The number of collaborating infrastructure sectors is equal to that specified in the alternative description , see Tab. SI.3 and Lienert et al.(2014b).
costcap	Annual cost per inhabitant in% of the mean taxable income [%]	+0.01-5	Covers costs for operation and maintenance of the water system, as well as expansion and re-investment, rehabilitation, and fees for import of water from the regional water supplier.	Calculated. Annual costs were calculated for 2010-2050 using unit cost estimates for expansion, rehabilitation, and operation and maintenance specified for following components: *Fees:* imported water fees (from regional water supplier), bottled water fees, water lorry delivery fee *Operation and maintenance of:* centralized water supply system, decentralized water storage (household tanks), decentralized firefighting tanks, point-of-entry (POE) treatment system, point-of-use (POU) treatment system, rainwater filters, decentralized tank chlorination. *Expansion of or reinvestment on supply system:* pipe rehabilitation, pipe network expansion, central water purification plant (WPP), central water reservoirs, central UV treatment, decentralized water storage (household tanks), decentralized firefighting tanks, POE systems, POU systems, rainwater filters, decentralized tank chlorination.
costchange	Mean annual (linear) increase of costs [%/a]	+0-5	Cost increases imply that additional financial resources have to be allocated.	Calculated. Derived from costcap using the annual linear increase of costs between 2010-2050.

SI3 Decision alternatives

Overview of decision alternatives

Table SI3.1: Technical specifications of decision alternatives. Other characteristics (organizational structure, sector cooperation, management, rehabilitation strategy, operation, and maintenance) are described in Lienert et al. (submitted). UV = ultra-violet disinfection; AOP = advanced oxidation process; GAC = granular activated carbon; POE = Point-of-entry treatment (e.g. in the cellar), POU = Point-of-use treatment (e.g. under the sink), O_3 = ozone, UF = ultrafiltration, RO = reverse osmosis.

No.	Name	Organization, cooperation, management	Rehabilitation, operation, and maintenance	Water supply and uses	Water sources	Water treatment technology
A1a	Centralized, privatization, high environmental protection	One private organization manages all sectors [(a)] and all municipalities [(b)] (also with entire region Zürich Oberland).	The rehabilitation strategy foresees 2% annual replacement by pipe condition. Extensive operation and maintenance in underground service galleries; average inspection.	Water is centrally treated and supplied for potable, household, and firefighting use. Dimensioning as usual.	2010 amounts from springs and groundwater wells, all the rest from regional water supplier (purified lake water).	Groundwater disinfection with UV; lake water treatment as today (multi-step treatment), but with AOP+GAC instead of current O_3-GAC.
A1b	Centralized, IKA	As A1a, but intercommunal agency (IKA) manages the infrastructure, not a contractor.	As A1a	As A1a	As A1a	As A1a
A2	Centralized IKA, rain stored	As A1b, but constant budget, 100% self-financed.	Rehabilitation is according to condition (1% annual replacement). The most economical pipe laying technique is used. Moderate operation, maintenance, and inspection.	Water is centrally treated and supplied for potable, household, and firefighting use. Dimensioning is on maximum hourly demand of households, further volumes for firefighting are held in decentralized underground firefighting water (FFW) tanks.	As A1a; rainwater is used as far as possible for filling firefighting water tanks.	Groundwater disinfection with UV; lake water treatment as today (multi-step treatment), but with AOP+GAC instead of current O3-GAC.
A3	Fully decentralized	All sectors [(a)] and communities [(b)] work separately. Main responsibility, also concerning funding, is with the consumers (households), who are well-informed. The services are contracted to external.	Only repairs, but no rehabilitation is undertaken, and only upon urgent need for action. Moderate operation and maintenance; little inspection.	Potable water for drinking and cooking from the supermarket, household water is treated in the households and delivered by water lorries. Fire-fighting water volumes are kept within the household water tank. Dimensioning as usual.	Household and firefighting water: as far as possible rain water from the roof and recycled grey water; drinking water from the supermarket. If household and firefighting water are not enough, people buy the necessary amounts individually from local and regional water suppliers.	Untreated groundwater from the case study area, lake water from regional water supplier (multi-step treatment with O_3-GAC). Water from the supermarket is purified spring or groundwater. Rainwater (after coarse filtration), grey water (after advanced treatment), and water delivered by lorries is stored in a tank and purified by a POE treatment module (GAC+UF) before use.

No.	Name	Organization, cooperation, management	Rehabilitation, operation, and maintenance	Water supply and uses	Water sources	Water treatment technology
A4	Decaying centralized infrastructure, decentralized outskirts	Water infrastructures are managed by a mix of municipalities, cooperatives, and households, and separately from other sectors [(a)]. Outside the urban area of 2010, private consumers are responsible. Specialized services are contracted to external companies.	As A3, but operation and maintenance is minimal.	Water for all purposes is centrally supplied in the area of 2010 (drinking water quality not ensured). In the outskirts, water is supplied by lorries once per week.	2010 amounts, if not enough, more water from regional water supplier (lake water).	No groundwater treatment for centralized supply, lake water treated equivalent to today's treatment. Households have their own POU drinking water treatment (GAC-RO filter)
A5	Decaying infrastructure everywhere	Most infrastructure services are in the responsibility of the customers (households), who are well-informed. Services are contracted to external organizations.	Measures are only undertaken upon urgent need for action, operation and maintenance are minimal (as A4), and no inspection at all.	No centralized water supply, no more pipes are built. Consumers operate tanks, which are intermittently recharged by a private delivery service with hygienically safe water (lorries). Fire-fighting volumes are stored in separate tanks.	All water is abstracted from springs and groundwater wells in the region.	In-house hygienization of tank water (chlorination).
A6	Maximal collaboration, centralized	There is maximal cooperation; the case study communities [(b)] and Oetwil am See are organized in a cooperative. This service provider combines water and wastewater services with telecommunication, electricity, gas, and road services [(a)].	Rehabilitation is done according to condition (1% annual replacement). Repair and replacement are done in trench. Their operation, maintenance, and inspection are moderate.	Centralized supply of drinking and household water. Dimensioning is on the maximum hourly demand of households, further volumes for firefighting are held in underground firefighting water tanks.	Withdrawal from sources and wells extended with rainwater so that only 10% of supply origins from the regional water supplier (lake ZH water). As much rainwater as possible is used for clothes washing and toilet flushing.	Lake water is treated (current multi-step with O_3-GAC), groundwater is not. Rainwater is coarsely filtrated at the inflow to the rainwater tank.
A7	Mixed responsibility, fully decentralized with onsite treatment	Public water supply and wastewater services are combined within one cooperative for all four case study communities [(b)]; no collaboration between different infrastructure services [(a)]. Private owners are responsible for treatment facilities on private grounds.	Rehabilitation is done according to prioritization. No rehabilitation of centralized system. Their operation, maintenance, and inspection are moderate.	Rainwater is reused in the households as far as possible. Further water will only be delivered by the municipality (lorries) upon special demand or in longer dry periods. Firefighting is provided by firefighting tanks (shared between neighboring lots).	2010 amounts from sources and wells in the region, all the rest from regional water supplier (lake water). As much rainwater as possible is used. The water demand is reduced through the use of urine diversion toilets.	POE treatment (GAC+UF) of all incoming water.
A8a	Status quo with storm water	The communities [(b)] remain responsible for a single, integrated wastewater and	Rehabilitation is done according to prioritization (1% annual replacement by	Water is centrally treated and supplied to be used as drinking, household, and firefighting water.	2010 amounts for centralized system, if not enough, more water is	Groundwater is disinfected (UV treatment), lake water receives a multi-step treatment as today,

No.	Name	Organization, cooperation, management	Rehabilitation, operation, and maintenance	Water supply and uses	Water sources	Water treatment technology
	retention	drinking water sector that jointly operate the water infrastructures, with some services contracted out to private enterprises.	condition and criticality). Renovation is trenchless. Their operation, maintenance, and inspection is moderate.		imported from the regional water supplier.	including O_3-GAC.
A8b	Status quo technical variant	As A8a	As A8a	Water is centrally treated and supplied to be used as drinking, household, and firefighting water. Newly developed housing areas are dimensioned on 30 m3/h fire flows –similar to 'self-cleaning networks'(Vreeburg et al., 2009).	As A8a	As A8a
A9	Centralized, privatization, minimal maintenance	The water infrastructures are fully contracted out, and all sectors work separately [(a)]. Private consumers choose their contracting provider.	Measures are only undertaken upon urgent need for action; only repair is done, and in trench. Operation and maintenance are minimal, with little inspection.	Centralized supply of drinking, household, and firefighting water, but dimensioning is the on maximum hourly demand of households. Further volumes for firefighting are held in underground fire-fighting water tanks.	As A8a	As A8a

[a] With all sectors we mean transportation, gas supply, energy supply, district heating, telecommunication, as well as water supply and wastewater disposal.

[b] The four communities are: Mönchaltorf, Gossau, Grüningen, Egg.

Prediction of attribute levels for alternatives

Table SI3.2: Predictions of the attributes (Tab. SI2.1) by alternative and scenario, stated as probability distributions. Explanation of abbreviations: A1a – A9…alternatives; see Table SI3.1 for a description; Status quo, Boom, Doom, Quality of life are the four socio-demographic future scenarios; DW… drinking water; HW… household water; FFW…firefighting water; β(x,y)…beta distribution with shape1 = x, shape2= y; N(x,y)…normal distribution with μ= x, σ= y; LN(x,y)…lognormal distribution with μ = x, σ= y; LOG(x,y)…logistic distribution with location = x, scale= y; U(x,y)…uniform distribution with min = x, max= y; TN(x,y [a,b])…truncated normal distribution with μ = x, σ = y and truncation at min= a, max = b.

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
Realization of the rehabilitation demand [%] (rehab)											
Status quo	β(9.0375, 4.0951)	β(9.0375, 4.0951)	β(19.0754, 8.9788)	U(0,0)	U(0,0)	U(0,0)	β(19.0754 8.9788	U(0,0)	N(0.0438, 0.0162)	N(0.0438, 0.0162)	U(0,0)
Boom	N(0.2486, 0.0814)	N(0.2486, 0.0814)	N(0.2027, 0.0744)	U(0,0)	U(0,0)	U(0,0)	N(0.2027, 0.0744)	U(0,0)	β(9.7487, 110.0828)	β(9.7487, 110.0828)	U(0,0)
Doom	β(9.0375, 4.0951)	β(9.0375, 4.0951)	β(19.0754, 8.9788)	U(0,0)	U(0,0)	U(0,0)	β(19.0754, 8.9788)	U(0,0)	N(0.0438, 0.0162)	N(0.0438, 0.0162)	U(0,0)
Quality of life	N(0.5692, 0.1517)	N(0.5692, 0.1517)	N(0.5212, 0.1261)	U(0,0)	U(0,0)	U(0,0)	N(0.5212, 0.1261)	U(0,0)	LOG(0.074, 0.0088)	LOG(0.074,0.0088)	U(0,0)
Flexibility of technical extension or deconstruction of infrastructure [%] (adapt)											
Status quo	N(35,7.65)	N(40,10.2)	N(20,10.2)	N(85,7.65)	N(62.5,6.38)	N(62.5,6.38)	N(55,7.65)	N(65,7.65)	N(35,7.65)	N(35,7.65)	N(30,10.2)
Boom	N(35,7.65)	N(40,10.2)	N(20,10.2)	N(85,7.65)	N(62.5,6.38)	N(62.5,6.38)	N(55,7.65)	N(65,7.65)	N(35,7.65)	N(35,7.65)	N(30,10.2)
Doom	N(35,7.65)	N(40,10.2)	N(20,10.2)	N(85,7.65)	N(62.5,6.38)	N(62.5,6.38)	N(55,7.65)	N(65,7.65)	N(35,7.65)	N(35,7.65)	N(30,10.2)
Quality of life	N(35,7.65)	N(40,10.2)	N(20,10.2)	N(85,7.65)	N(62.5,6.38)	N(62.5,6.38)	N(55,7.65)	N(65,7.65)	N(35,7.65)	N(35,7.65)	N(30,10.2)
% Utilization of groundwater recharge [%] (gwhh)											
Status quo	N(6.45,1.08)	N(6.45,1.08)	N(6.45,1.08)	N(5.32,0.89)	N(6.45,1.08)	N(11,1.84)	N(8.49,1.42)	N(6.45,1.08)	N(6.45,1.08)	N(6.45,1.08)	N(6.45,1.08)
Boom	N(7.51,1.25)	N(7.51,1.25)	N(7.51,1.25)	N(81.66,13.64)	N(7.51,1.25)	N(134.69, 22.49)	N(118.96, 19.87)	N(7.51,1.25)	N(7.51,1.25)	N(7.51,1.25)	N(7.51,1.25)
Doom	N(6.45,1.08)	N(6.45,1.08)	N(6.45,1.08)	N(3.57,0.6)	N(6.45,1.08)	N(10.55,1.76)	N(7.84,1.31)	N(6.45,1.08)	N(6.45,1.08)	N(6.45,1.08)	N(6.45,1.08)
Quality of life	N(6.50,1.09)	N(6.5,1.09)	N(6.5,1.09)	N(6.37,1.06)	N(6.37,1.06)	N(12.71,2.12)	N(9.93,1.66)	N(6.5,1.09)	N(6.5,1.09)	N(6.5,1.09)	N(6.5,1.09)
Net energy consumption for water treatment and transport [kWh/m3] (econs)											
Status quo	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.0777, 0.0194)	N(0.4,0.1)	N(0.3649, 0.0912)	N(0.55, 0.1375)	N(0.185, 0.0462)	N(0.67, 0.1675)	N(0.67, 0.1675)	N(0.67, 0.1675)
Boom	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.119, 0.0298)	N(0.2996, 0.0749)	N(0.3649, 0.0912)	N(0.55, 0.1375)	N(0.2654, 0.0664)	N(0.67, 0.1675)	N(0.67, 0.1675)	N(0.67, 0.1675)
Doom	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.0898, 0.0225)	N(0.4,0.1)	N(0.3649, 0.0912)	N(0.55, 0.1375)	N(0.2148, 0.0537)	N(0.67, 0.1675)	N(0.67, 0.1675)	N(0.67, 0.1675)
Quality of life	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.713, 0.1783)	N(0.0778, 0.0194)	N(0.4,0.1)	N(0.3649, 0.0912)	N(0.55, 0.1375)	N(0.1797, 0.0449)	N(0.67, 0.1675)	N(0.67, 0.1675)	N(0.67, 0.1675)
DW: Days per year with water quantity limitations [d/a] (vol_dw)											

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
Status quo	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
Boom	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
Doom	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
Quality of life	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
HW: Days per year with water quantity limitations [d/a] (vol_hw)											
Status quo	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	NU0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
Boom	U(0,0)	U(0,0)	U(0,0)	U(0,0)	N(18.66, 0.9006)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
Doom	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
Quality of life	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)	U(0,0)
FFW: Available water for firefighting in new housing areas [L/min] (vol_ffw)											
Status quo	N(1766.968, 442)	N(1766.968, 442)	N(1310.211, 328)	N(1726.288, 432)	N(1766.968, 442)	N(1838.676, 460)	N(1310.211, 328)	N(1838.676, 460)	N(1766.968, 442)	N(1766.968, 442)	N(1310.211, 32)8
Boom	N(3600,900)	N(3600,900)	N(3600,900)	N(2902.984, 726)	N(3600,900)	N(3600,900)	N(3600,900)	N(3600,900)	N(3600,900)	N(3600,900)	N(3600, 900)
Doom	N(1854.309, 464)	N(1854.309, 464)	N(1497.555, 375)	N(1791.37, 448)	N(1854.309,4 64)	N(1960.12, 491)	N(1497.555, 375)	N(1960.12, 491)	N(1854.309, 464)	N(1854.309, 464)	N(1497.555,3 75)
Quality of life	N(1766.968, 442)	N(1766.968, 442)	N(1310.211, 328)	N(1726.288, 432)	N(1766.968, 442)	N(1838.676, 460)	N(1310.211, 328)	N(1838.676, 460)	N(1766.968, 442)	N(1766.968, 442)	N(1310.211,3 28)
DW: System reliability (in interviews: "criticality") [-] (reliab_dw)											
Status quo	LN(-5.2162, 0.2991)	LN(-5.2162, 0.2991)	LN(-5.1793, 0.3056)	U(0.98,1)	N(0.0827, 0.0161)	N(0.175, 0.0375)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.2198, 0.3378)	LN(-4.2198, 0.3378)	LN(-4.0617, 0.3748)
Boom	β(2.5936, 694.7973)	β(2.5936, 694.7973)	β(2.7087, 689.5533)	U(0.98,1)	U(0.98,1)	N(0.175, 0.0375)	β(2.7087, 689.5533)	N(0.065, 0.0175)	β(2.8013, 680.5096)	β(2.8013, 680.5096)	β(3.0522, 653.4647)
Doom	LN(-5.2162, 0.2991)	LN(-5.2162, 0.2991)	LN(-5.1793, 0.3056)	U(0.98,1)	N(0.0827, 0.0161)	N(0.175, 0.0375)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.2198, 0.3378)	LN(-4.2198, 0.3378)	LN(-4.0617, 0.3748)
Quality of life	Beta(4.073, 688.1364)	Beta(4.073, 688.1364)	LN(-5.1757, 0.4138)	U(0.98,1)	N(0.0897, 0.0171)	N(0.175, 0.0375)	LN(-5.1757, 0.4138)	N(0.065, 0.0175)	LN(-4.7867, 0.3619)	LN(-4.7867, 0.3619)	LN(-4.5669, 0.3502)
HW: System reliability (in interviews: "criticality") [-] (reliab_hw)											
Status quo	LN(-5.2162, 0.2991)	LN(-5.2162, 0.2991)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.0617, 0.3748)	N(0.175, 0.0375)	LN(-5.1793, 0.3056)	N(0.175, 0.0375)	LN(-4.2198; 0.3378)	LN(-4.2198, 0.3378)	LN(-4.0617, 0.3748)
Boom	β(2.5936, 694.7973)	β(2.5936, 694.7973)	β(2.7087, 689.5533)	N(0.065, 0.0175)	N(0.0878, 0.0163)	N(0.175, 0.0375)	β(2.7087, 689.5533)	N(0.175, 0.0375)	β(2.8013, 680.5096)	β(2.8013, 680.5096)	β(3.0522, 653.4647)
Doom	LN(-5.2162, 0.2991)	LN(-5.2162, 0.2991)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.0617, 0.3748)	N(0.175, 0.0375)	LN(-5.1793, 0.3056)	N(0.175, 0.0375)	LN(-4.2198, 0.3378)	LN(-4.2198, 0.3378)	LN(-4.0617, 0.3748)
Quality of life	β(4.073, 688.1364)	β(4.073, 688.1364)	LN(-5.1757, 0.4138)	N(0.065, 0.0175)	N(0.055, 0.0107)	N(0.175, 0.0375)	LN(-5.1757, 0.4138)	N(0.175, 0.0375)	LN(-4.7867, 0.3619)	LN(-4.7867, 0.3619)	LN(-4.5669, 0.3502)
FFW: System reliability (in interviews: "criticality") [-] (reliab_ffw)											

	A1a	**A1b**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8a**	**A8b**	**A9**
Status quo	LN(-5.2162, 0.2991)	LN(-5.2162, 0.2991)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.0617, 0.3748)	N(0.175, 0.0375)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.2198, 0.3378)	LN(-4.2198, 0.3378	LN(-4.0617, 0.3748)
Boom	β(2.5936, 694.7973)	β(2.5936, 694.7973)	β(2.7087, 689.5533)	N(0.065, 0.0175)	N(0.0638, 0.0118)	N(0.175, 0.0375)	β(2.7087, 689.5533)	N(0.065, 0.0175)	β(2.8013, 680.5096)	β(2.8013 680.5096	β(3.0522, 653.4647)
Doom	LN(-5.2162, 0.2991)	LN(-5.2162, 0.2991)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.0617, 0.3748)	N(0.175, 0.0375)	LN(-5.1793, 0.3056)	N(0.065, 0.0175)	LN(-4.2198, 0.3378)	LN(-4.2198, 0.3378)	LN(-4.0617, 0.3748)
Quality of life	β(4.073, 688.1364)	β(4.073, 688.1364)	LN(-5.1757, 0.4138)	N(0.065, 0.0175)	LN(-3.2535, 0.2143)	N(0.175, 0.0375)	LN(-5.1757, 0.4138)	N(0.065, 0.0175)	LN(-4.7867, 0.3619)	LN(-4.7867, 0.3619)	LN(-4.5669, 0.3502)
DW: Days per year with esthetic impairment such as taste, smell, etc.[d/a] (aes_dw)											
Status quo	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(1,0.51)	N(1,0.51)	N(20, 5.1)	N(5,2.55)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(10,5.1)
Boom	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(1,0.51)	N(1,0.51)	N(20, 5.1)	N(5,2.55)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(15,7.65)
Doom	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(1,0.51)	N(1,0.51)	N(20, 5.1)	N(5,2.55)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(10,5.1)
Quality of life	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(1,0.51)	N(1,0.51)	N(20, 5.1)	N(5,2.55)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(10,5.1)
HW: Days per year with esthetic impairment such as taste, smell, etc.[d/a] (aes_hw)											
Status quo	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(55,22.96)	N(75,12.76)	N(20, 5.1)	N(10,5.1)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(10,5.1)
Boom	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(55,22.96)	N(75,12.76)	N(20, 5.1)	N(10,5.1)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(15,7.65)
Doom	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(55,22.96)	N(75,12.76)	N(20, 5.1)	N(10,5.1)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(10,5.1)
Quality of life	N(5,2.55)	N(5,2.55)	N(5,2.55)	N(55,22.96)	N(75,12.76)	N(20, 5.1)	N(10,5.1)	N(27.5,11.48)	N(5,2.55)	N(5,2.55)	N(10,5.1)
DW: Days per year with hygienic concerns (hygiene indicators) [d/a] (faecal_dw)											
Status quo	N(2.5,1.28)	N(2.5,1.28)	N(2.5,1.28)	U(0,0)	U(0,0)	N(1,0.51)	U(0,0)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
Boom	N(1,0.51)	N(1,0.51)	N(1,0.51)	U(0,0)	U(0,0)	N(1,0.51)	U(0,0)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
Doom	N(2.5,1.28)	N(2.5,1.28)	N(2.5,1.28)	U(0,0)	U(0,0)	N(1,0.51)	U(0,0)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
Quality of life	N(2.5,1.28)	N(2.5,1.28)	N(2.5,1.28)	U(0,0)	U(0,0)	N(1,0.51)	U(0,0)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
HW: Days per year with hygienic concerns (hygiene indicators) [d/a] (faecal_hw)											
Status quo	N(2.5,1.28)	N(2.5,1.28)	N(2.5,1.28)	U(0,0)	N(20, 5.1)	N(1,0.51)	N(5,2.55)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
Boom	N(1,0.51)	N(1,0.51)	N(1,0.51)	U(0,0)	N(20, 5.1)	N(1,0.51)	N(5,2.55)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
Doom	N(2.5,1.28)	N(2.5,1.28)	N(2.5,1.28)	U(0,0)	N(20, 5.1)	N(1,0.51)	N(5,2.55)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
Quality of life	N(2.5,1.28)	N(2.5,1.28)	N(2.5,1.28)	U(0,0)	N(20, 5.1)	N(1,0.51)	N(5,2.55)	U(0,0)	N(2.5,1.28)	N(2.5,1.28)	N(5,2.55)
DW: Changes in total cell count as indicator of bacterial regrowth [log units] (cells_dw)											
Status quo	U(0,0)	U(0,0)	U(0,0)	N(0.15,0.08)	N(-0.5,0.26)	N(-1.5,0.26)	N(0.14,0.07)	N(0.34,0.07)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)
Boom	N(0.1,0.05)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)	N(-0.85,0.59)	N(-1.5,0.26)	N(0.14,0.07)	N(0.34,0.07)	N(0.15,0.08)	N(0.15,0.08)	N(0.15,0.08)

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
Doom	N(0.1,0.05)	N(0.1,0.05)	U(0,0)	N(0.15,0.08)	N(-0.5,0.26)	N(-1.5,0.26)	N(0.14,0.07)	N(0.34,0.07)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)
Quality of life	U(0,0)	U(0,0)	U(0,0)	N(0.15,0.08)	N(-0.5,0.26)	N(-1.5,0.26)	N(0.14,0.07)	N(0.34,0.07)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)
HW: Changes in total cell count as indicator of bacterial regrowth [log units] (cells_hw)											
Status quo	U(0,0)	U(0,0)	N(0.1,0.05)	N(0.39,0.05)	N(0.35,0.18)	N(-1.5,0.26)	N(0.24,0.03)	N(0.34,0.07)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)
Boom	N(0.1,0.05)	N(0.1,0.05)	N(0.1,0.05)	N(0.39,0.05)	N(-0.65,0.69)	N(-1.5,0.26)	N(0.23,0.03)	N(0.34,0.07)	N(0.15,0.08)	N(0.15,0.08)	N(0.15,0.08)
Doom	N(0.1,0.05)	U(0,0)	N(0.1,0.05)	N(0.39,0.05)	N(0.35,0.18)	N(-1.5,0.26)	N(0.24,0.03)	N(0.34,0.07)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)
Quality of life	U(0,0)	U(0,0)	N(0.1,0.05)	N(0.39,0.05)	N(0.35,0.18)	N(-1.5,0.26)	N(0.24,0.03)	N(0.34,0.07)	N(0.1,0.05)	N(0.1,0.05)	N(0.15,0.08)
DW and HW: Inorganic substances (indicator: nitrate concentration) [mg/L] (no3_dw, no3_hw)											
Status quo	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)
Boom	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)
Doom	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)
Quality of life	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)	U(0,20)
DW and HW: Pesticides (sum of pesticide concentration) [µg/L] (pest_dw, pest_hw)											
Status quo	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)
Boom	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)
Doom	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)
Quality of life	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)	U(0,0.02)
DW and HW: Micropollutants (indicator: benzotriazole) [ng/L] (bta_dw, bta_hw)											
Status quo	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)
Boom	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)
Doom	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)
Quality of life	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)	U(0,150)
Score of the EFQM excellence model (European Foundation for Quality Management) [%] (efqm)											
Status quo	N(68, 6.63)	N(72,6.63)	N(69,4.59)	N(37,5.61)	N(39,7.65)	N(33,5.61)	N(65,2.55)	N(62,5.1)	N(63,2.55)	N(63,2.55)	N(46,8.16)
Boom	N(72,4.59)	N(72,6.63)	N(71,4.59)	N(39,5.61)	N(41,7.65)	N(35,5.61)	N(69,2.55))	N(60,5.1)	N(63,2.55)	N(63,2.55)	N(48,8.16)
Doom	N(67, 6.12)	N(70,6.63)	N(66,5.1)	N(35,5.61)	N(37,7.65)	N(31,5.61)	N(63,2.55)	N(64,5.1)	N(65,2.55)	N(65,2.55)	N(42,8.16)
Quality of life	N(72,4.59)	N(72,6.63)	N(71,4.59)	N(37,5.61)	N(39,7.65)	N(33,5.61)	N(65,2.55)	N(62,5.1)	N(63,2.55)	N(63,2.55)	N(46,8.16)
Degree (percent) of codetermination [%] (voice)											

	A1a	**A1b**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8a**	**A8b**	**A9**
Status quo	N(20,10.2)	N(40,10.2)	N(50,4.51)	N(80,10.2)	N(70,15.31)	N(80,10.2)	N(60,10.2)	N(75,12.76)	N(70,10.2)	N(70,10.2)	N(80,10.2)
Boom	N(20,10.2)	N(40,10.2)	N(50,4.51)	N(80,10.2)	N(70,15.31)	N(80,10.2)	N(60,10.2)	N(75,12.76)	N(70,10.2)	N(70,10.2)	N(80,10.2)
Doom	N(20,10.2)	N(40,10.2)	N(50,4.51)	N(80,10.2)	N(70,15.31)	N(80,10.2)	N(60,10.2)	N(75,12.76)	N(70,10.2)	N(70,10.2)	N(80,10.2)
Quality of life	N(20,10.2)	N(40,10.2)	N(50,4.51)	N(80,10.2)	N(70,15.31)	N(80,10.2)	N(60,10.2)	N(75,12.76)	N(70,10.2)	N(70,10.2)	N(80,10.2)
% of water coming from the Mönchaltorfer Aa region [%] (auton)											
Status quo	U(55.1981, 55.1981)	U(55.2, 55.2)	U(55.2, 55.2)	U(80.32, 80.32)	U(55.46, 55.46)	U(100,100)	U(90,90)	U(89.33, 89.33)	U(55.46, 55.46)	U(55.4571, 55.4571)	U(55.46, 55.46)
Boom	U(5.25, 5.25)	U(5.25, 5.25)	U(5.25, 5.25)	U(80.32, 80.32)	U(55.46, 55.46)	U(100,100)	U(90,90)	U(89.33, 89.33)	U(55.46, 55.46)	U(55.4571, 55.4571)	U(55.46, 55.46)
Doom	U(57.58, 57.58)	U(57.58, 57.58)	U(57.58, 57.58)	U(80.32, 80.32)	U(55.46, 55.46)	U(100,100)	U(90,90)	U(89.33, 89.33)	U(55.46, 55.46)	U(55.4571, 55.4571)	U(55.46, 55.46)
Quality of life	U(48.1738, 48.1738)	U(48.1738, 48.1738)	U(48.1738, 48.1738)	U(81.0792, 81.0792)	U(48.3998, 48.3998)	U(100,100)	U(90,90)	U(81.1719, 81.1719)	U(48.3998, 48.3998)	U(48.3998, 48.3998)	U(48.3998, 48.3998)
Necessary time investment for operation and maintenance by user [h/(inh.*a)] (time)											
Status quo	U(0,0)	U(0,0)	U(0.36, 0.36)	U(1.69, 1.69)	U(5,5)	U(8.04, 8.04)	U(0.36, 0.36)	U(1.69, 1.69)	U(0,0)	U(0,0)	U(0,0)
Boom	U(0,0)	U(0,0)	U(0.12, 0.12)	U(0.9, 0.9)	U(0,0)	U(4.94, 4.94)	U(0.12, 0.12)	U(0.9, 0.9)	U(0,0)	U(0,0)	U(0,0)
Doom	U(0,0)	U(0,0)	U(0.36, 0.36)	U(1.33, 1.33)	U(5,5)	U(9.65, 9.65)	U(0.36, 0.36)	U(1.69, 1.69)	U(0,0)	U(0,0)	U(0,0)
Quality of life	U(0,0)	U(0,0)	U(0.3326, 0.3326)	U(1.4917, 1.4917)	U(4.9064, 4.9064)	U(6.9569, 6.9569)	U(0.3326, 0.3326)	U(1.595, 1.595)	U(0,0)	U(0,0)	U(0,0)
Additional area demand on private property per end user [m2/inh] (area)											
Status quo	U(0,0)	U(0,0)	U(0,0)	U(7.35, 7.35)	U(0.25, 0.25)	U(5.63, 5.63)	U(6.78, 6.78)	U(7.09, 7.09)	U(0,0)	U(0,0)	U(0,0)
Boom	U(0,0)	U(0,0)	U(0.57, 0.57)	U(7.35, 7.35)	U(0.25, 0.25)	U(5.63, 5.63)	U(6.78, 6.78)	U(7.09, 7.09)	U(0,0)	U(0,0)	U(0,0)
Doom	U(0,0)	U(0,0)	U(0,0)	U(7.35, 7.35)	U(0.25, 0.25)	U(5.63, 5.63)	U(6.78, 6.78)	U(7.09, 7.09)	U(0,0)	U(0,0)	U(0,0)
Quality of life	U(0,0)	U(0,0)	U(0.3545, 0.3545)	U(7.1232, 7.1232)	U(0.2453, 0.2453)	U(5.4039, 5.4039)	U(6.515, 6.515)	U(6.7414, 6.7414)	U(0,0)	U(0,0)	U(0.3545, 0.3545)
Number of infrastructure sectors that collaborate in planning and construction [-] (collab)											
Status quo	U(6,6)	U(6,6)	U(6,6)	U(1,1)	U(1,1)	U(2,2)	U(6,6)	U(6,6)	U(2,2)	U(2,2)	U(1,1)
Boom	U(6,6)	U(6,6)	U(6,6)	U(1,1)	U(1,1)	U(2,2)	U(6,6)	U(6,6)	U(2,2)	U(2,2)	U(1,1)
Doom	U(6,6)	U(6,6)	U(6,6)	U(1,1)	U(1,1)	U(2,2)	U(6,6)	U(6,6)	U(2,2)	U(2,2)	U(1,1)
Quality of life	U(6,6)	U(6,6)	U(6,6)	U(1,1)	U(1,1)	U(2,2)	U(6,6)	U(6,6)	U(2,2)	U(2,2)	U(1,1)
Annual cost per person in% of the mean taxable income [%] (costcap)											

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
Status quo	LN(-5.1776, 0.1232)	LN(-5.1776, 0.1232)	TN(0.0039, 0.0006)[0.002, 0.007]	LN(-4.2529, 0.2835)	LN(-5.6495, 0.1676)	LN(-5.0688, 0.3677)	TN(0.0039, 0.0006)[0.002, 0.006]	LN(-4.7923, 0.2947)	LN(-5.5707, 0.1603)	LN(-5.5707, 0.1603)	β(25.88, 8599.462)
Boom	U(0.0346, 0.0565)	U(0.0346, 0.0565)	U(0.02, 0.04)	U(0.0018, 0.0225)	U(0.0015, 0.021)	U(0.0007, 0.0052)	U(0.016, 0.0365)	β(10.9985 5798.49	U(0.0101, 0.0432)	U(0.0085, 0.0359)	U(0.0147, 0.0327)
Doom	LN(-4.3689, 0.1219)	LN(-4.3689, 0.1219)	LN(-4.745, 0.1434)	TN(0.035, 0.0092)[0,0.08]	LN(-4.8506, 0.1726)	LN(-4.2149, 0.3446)	TN(0.0087, 0.0013)[0.004, 0.014]	TN(0.02, 0.0127) [0,0.2]	TN(0.0085, 0.0014)[0.004, 0.014]	TN(0.0085, 0.0014)[0.004, 0.014]	TN(0.0066, 0.0012)[0.002, 0.012]
Quality of life	U(0.0088, 0.0147)	U(0.0088, 0.0147)	U(0.0042, 0.0091)	β(12.4288, 1453.01)	U(0.004, 0.009)	LN(-5.6628, 0.3674)	U(0.0043, 0.0093)	LN(-5.3033, 0.2926)	U(0.003, 0.0102)	U(0.003, 0.0102)	U(0.0034, 0.0075)
Mean annual (linear) increase of costs [%/a] (costchange)											
Status quo	N(0.0062, 0.0003)	N(0.0062, 0.0003)	N(0.0043, 0.0002)	N(0.0043, 0.0002)	N(0.0038, 0.0002)	N(0.0074, 0.0004)	N(0.0043, 0.0002)	N(0.0094, 0.0005)	N(0.0042, 0.0002)	N(0.0042, 0.0002)	N(0.0032, 0.0001)
Boom	N(0.0216, 0.017)	N(0.0216, 0.017)	N(0.0138, 0.009)	N(0.0297, 0.0138)	N(0.0242, 0.0112)	N(0.0042, 0.002)	N(0.0154, 0.0085)	N(0.0094, 0.0005)	N(0.0136, 0.0093)	N(0.012, 0.0076)	N(0.0128, 0.0086)
Doom	N(0.0095, 0.0018)	N(0.0095, 0.0018)	N(0.0066, 0.0012)	N(0.0264, 0.0047)	N(0.0059, 0.0011)	N(0.0118, 0.0021)	N(0.0065, 0.0012)	N(0.0151, 0.0027)	N(0.0063, 0.0012)	N(0.0063, 0.0012)	N(0.0049, 0.0009)
Quality of life	N(0.0096, 0.0014)	N(0.0096, 0.0014)	N(0.0061, 0.0006)	N(0.013, 0.0031)	N(0.0059, 0.0007)	N(0.0057, 0.0014)	N(0.006, 0.0005)	N(0.008, 0.0019)	N(0.0059, 0.0006)	N(0.0059, 0.0006)	N(0.0049, 0.0005)

Attribute predictions for the alternatives (sample size n= 10'000)

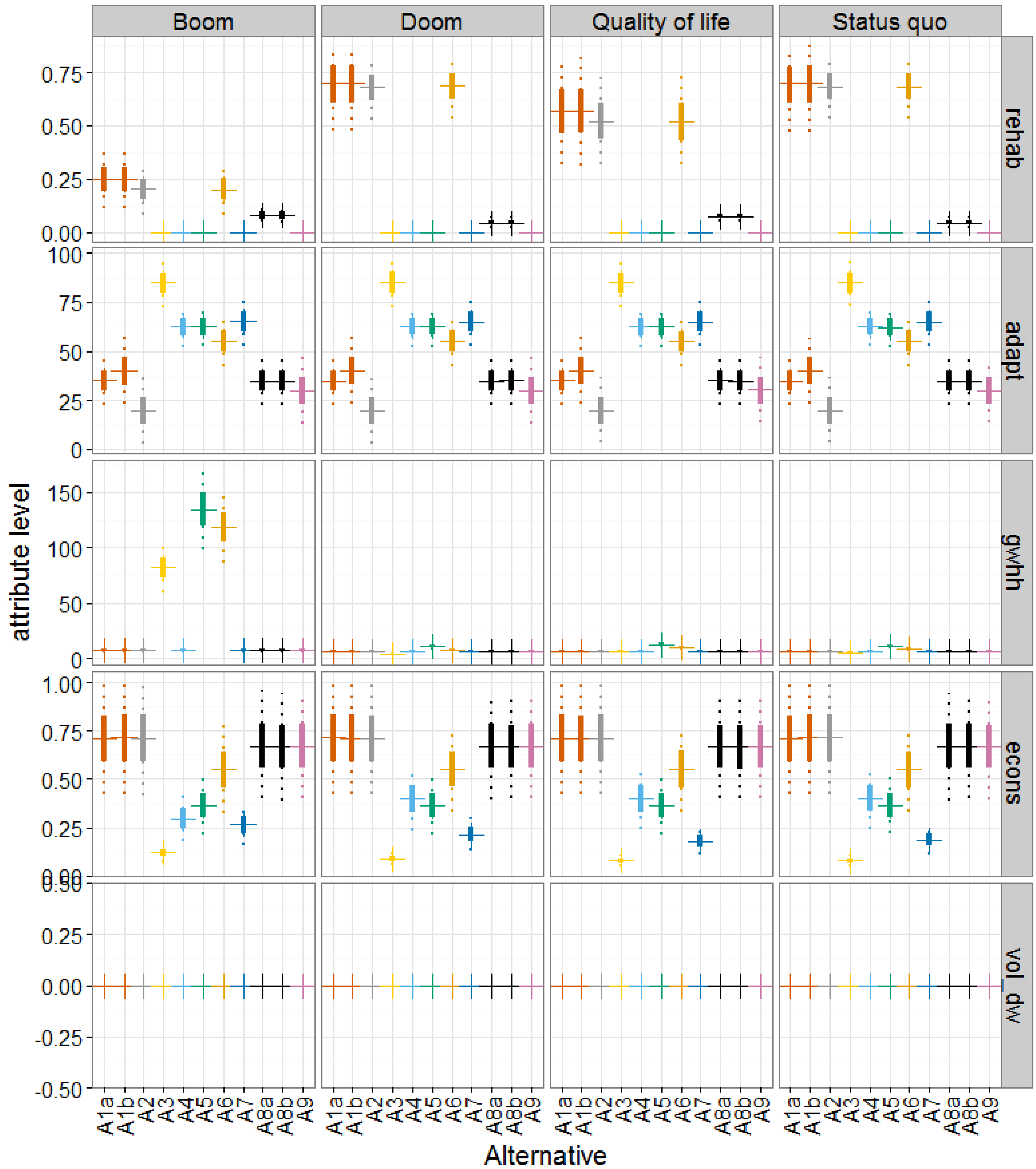


Figure SI3.1a: Distribution of attribute levels by alternative (A1a to A9; see Tab. SI.4) for four future scenarios (Boom, Doom, Quality of lifelife, Status quo). Labels on the right correspond to the short names of the attributes as listed in Table SI2.1. The predicted attribute levels and their uncertainty are given on the y-axis (see Tab. SI3.1 for attribute units and range). Thick, solid lines represent the 25 to 75% quantiles, dotted lines the 5 to 95% quantiles, the horizontal dash/cross the median.

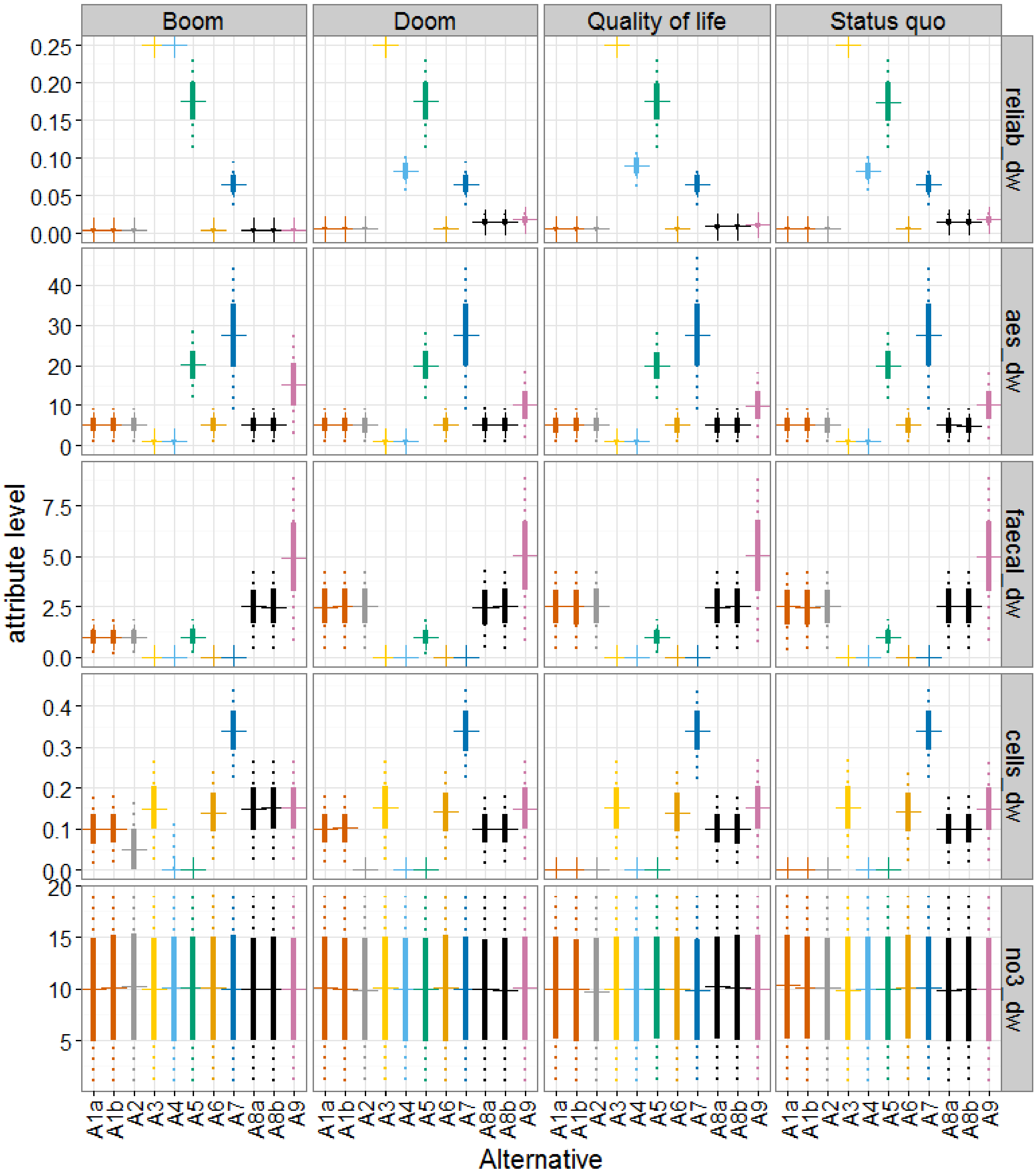


Figure SI3.1b For descriptions see Fig. SI3.1a.

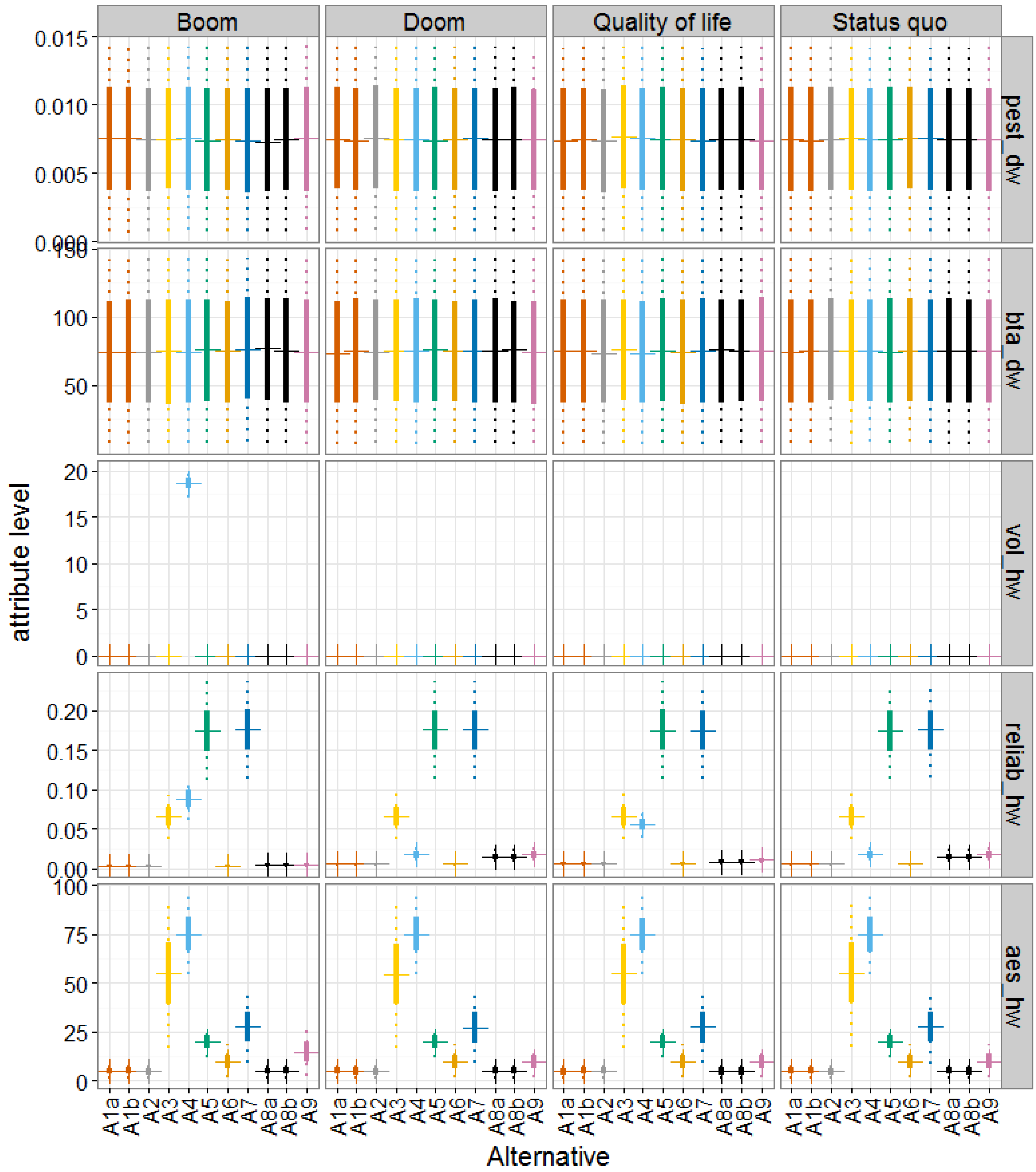


Figure SI3.1c For descriptions see Fig. SI3.1a.

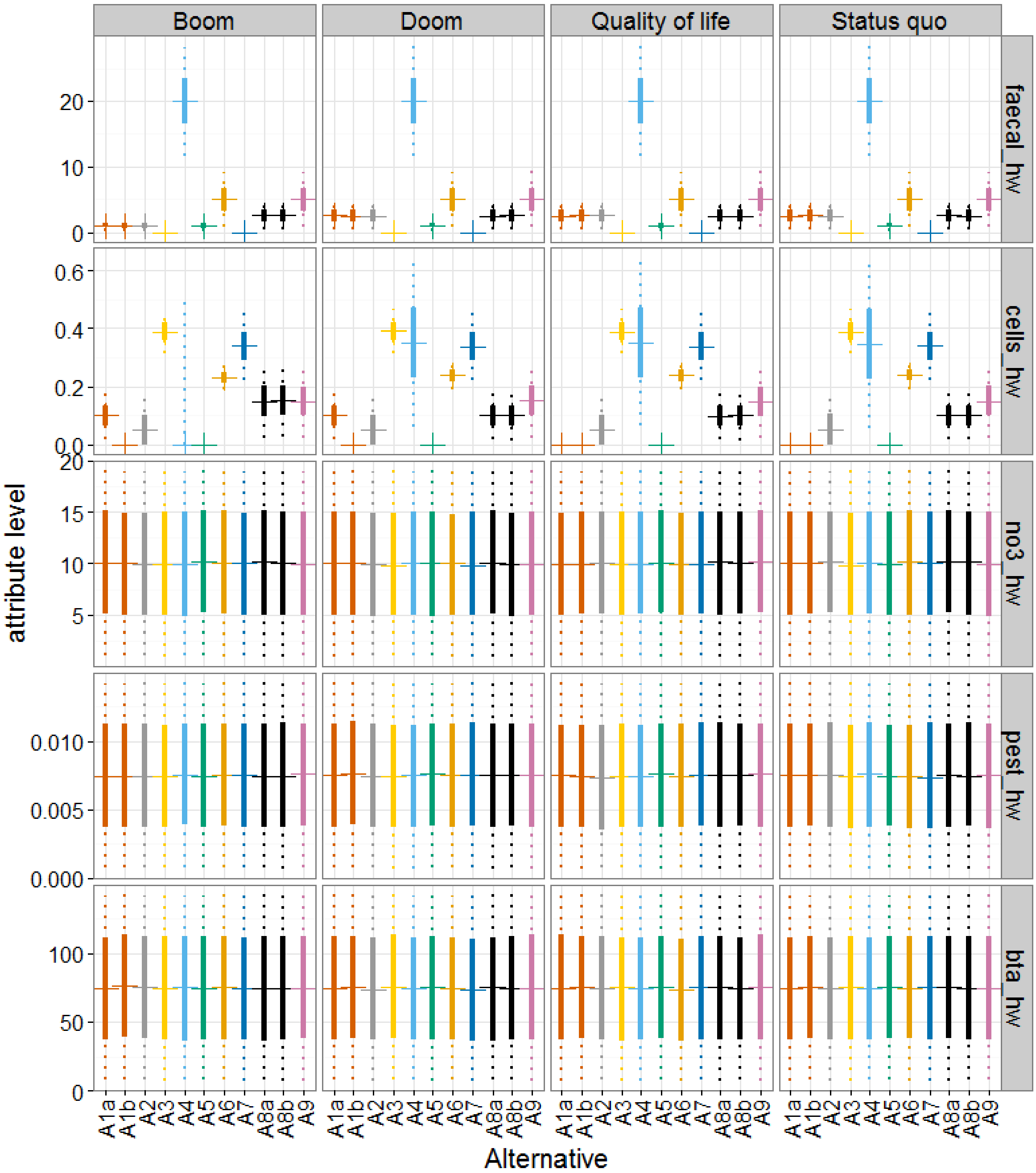


Figure SI3.1d For descriptions see Fig. SI3.1a.

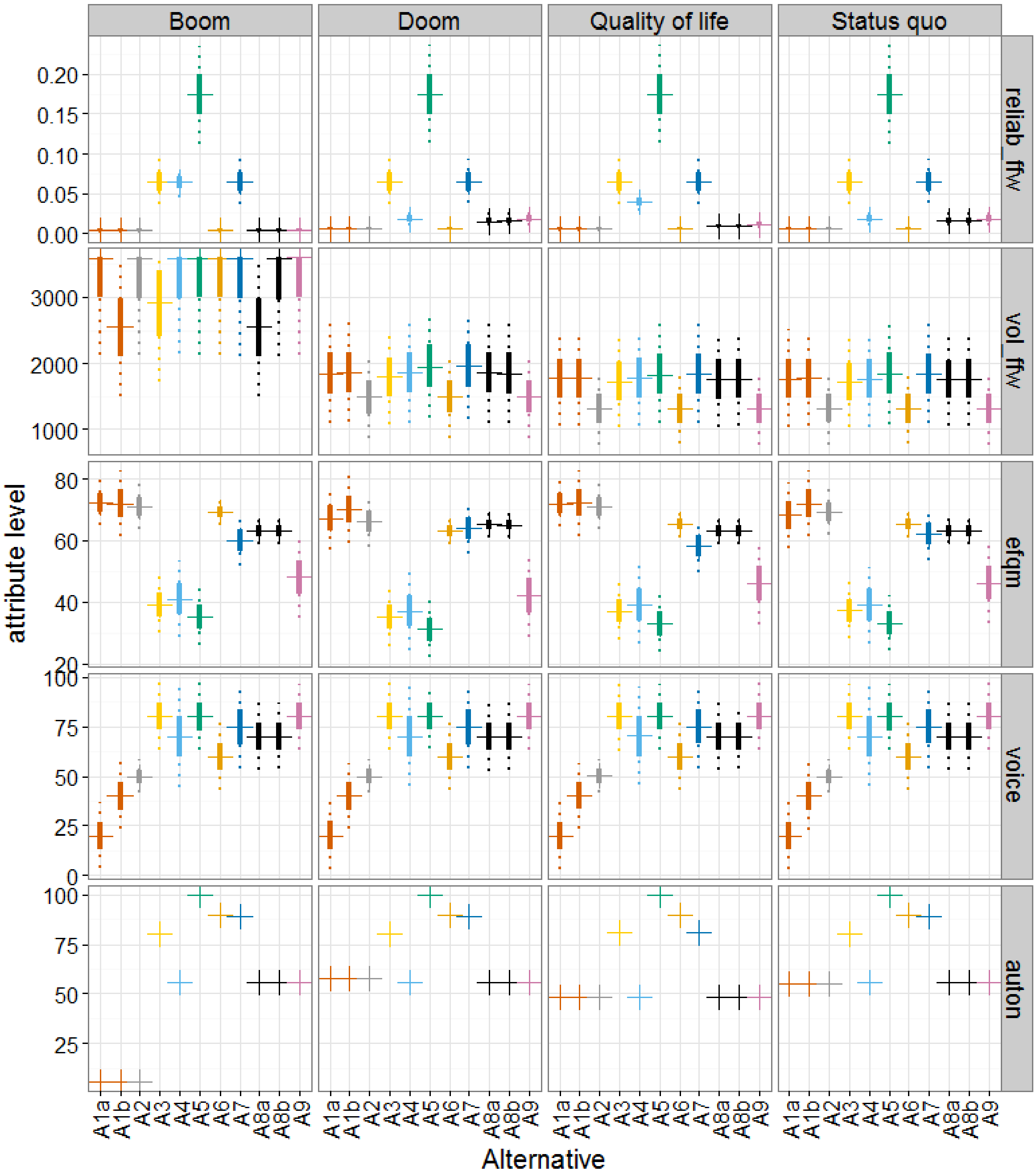
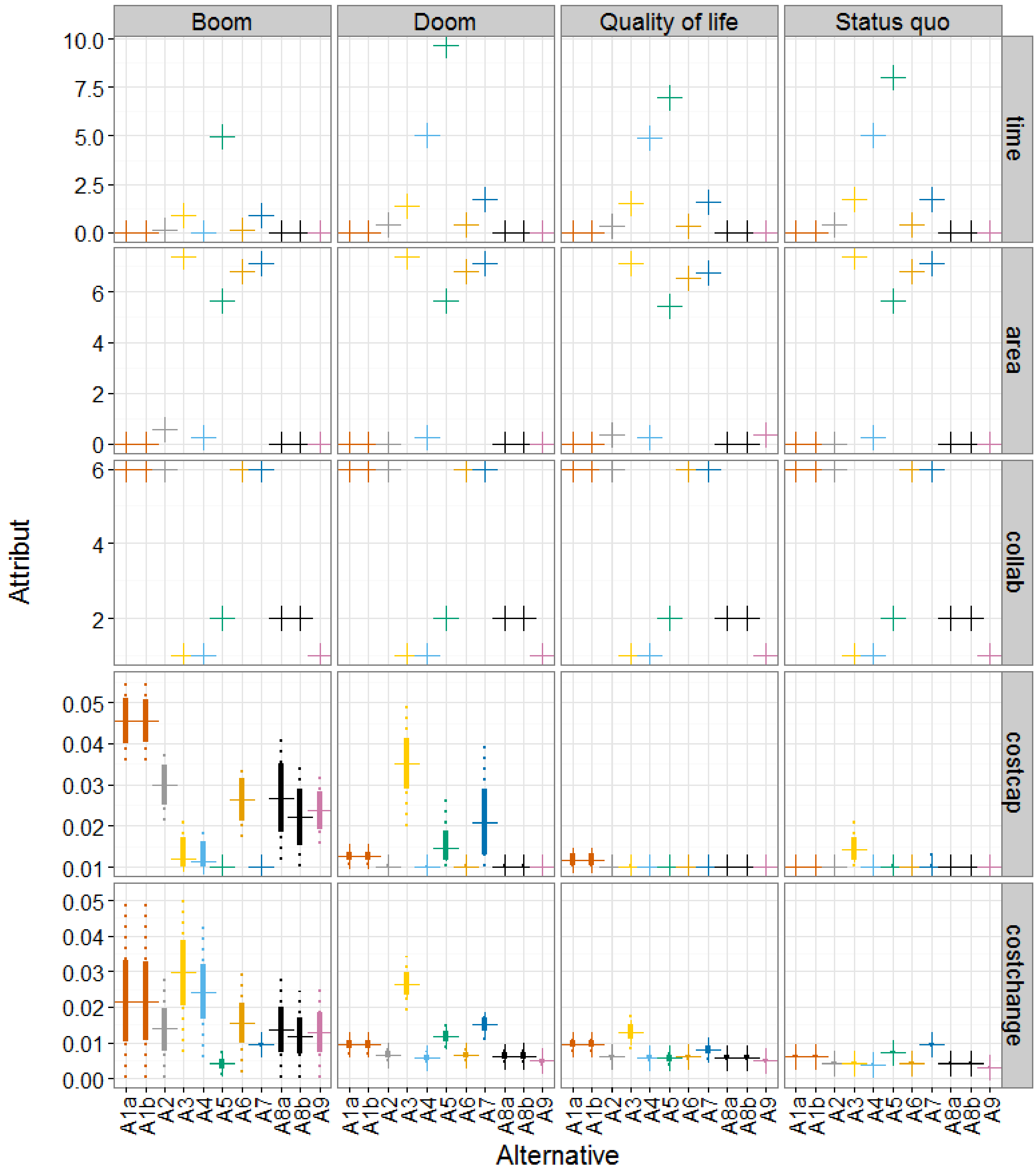


Figure SI3.1e For descriptions see Fig. SI3.1a.


Boom
Doom
Quality of life
Status quo
time
area
collab
costcap
costchange
Attribut
Alternative
A1a
A1b
A2
A3
A4
A5
A6
A7
A8a
A8b
A9
10.0
7.5
5.0
2.5
0.0
6
4
2
0
0.05
0.04
0.03
0.02
0.01
0.00

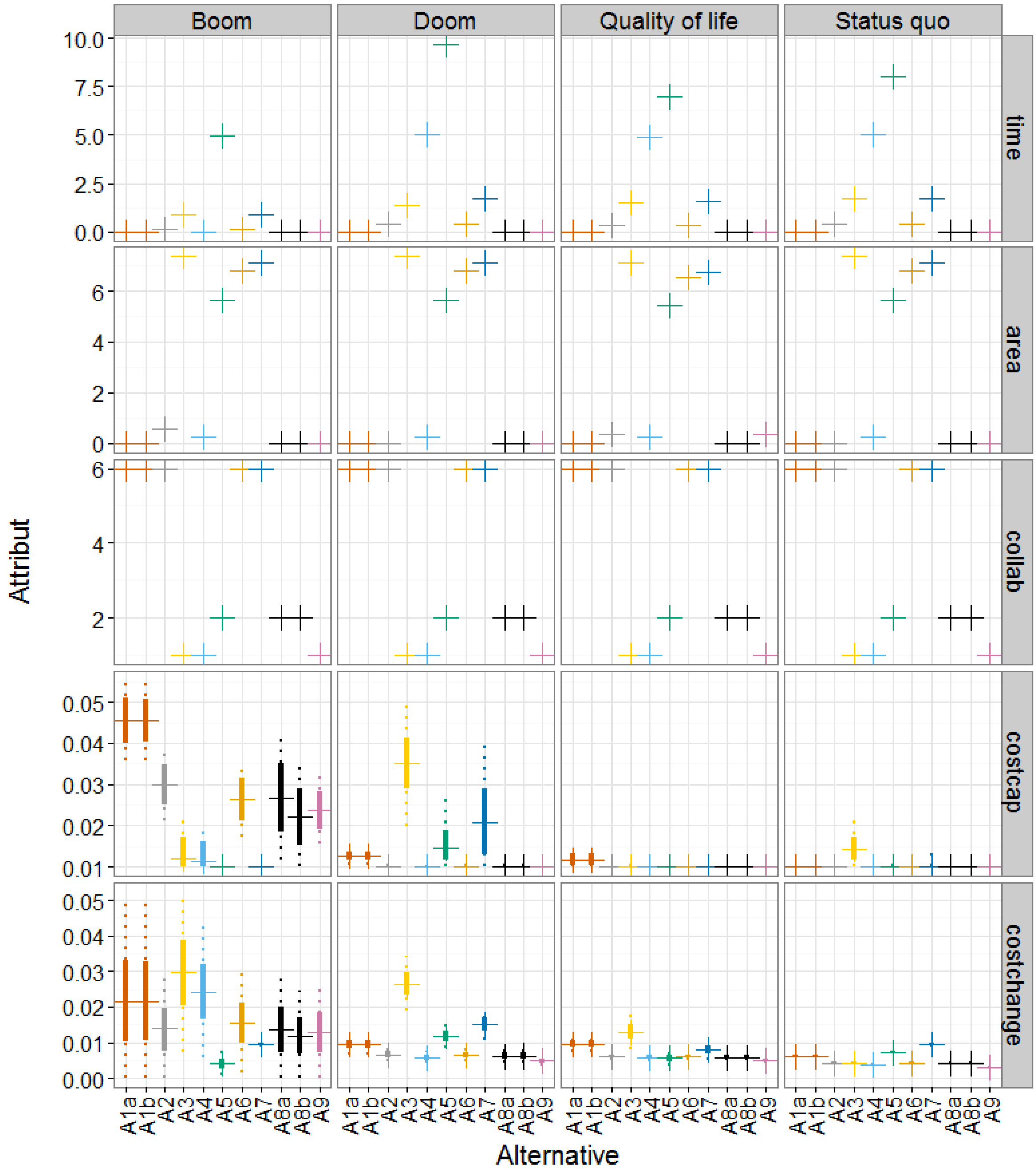


Figure SI3.1f For descriptions see Fig. SI3.1a.

SI4 Stakeholder preferences

Weights

Table SI4.1: Elicited weights from face-to-face interviews with ten stakeholders (see Tab. SI1.1). The order in the table follows our top-down elicitation procedure, starting with the five fundamental objectives at the highest level of the objectives hierarchy, and then moving systematically downwards in the hierarchy to the attribute level (see Tab. SI2.1). Objectives receiving zero weight (= irrelevant) are grey shaded. SH = stakeholder, dw = drinking water, hw= household water, ffw = firefighting water.

Objective	Weight no.	SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	over-all
		min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max
Intergenera-tional equity	w.1	0.13	0.16	0.10	0.26	0.25	0.19	0.13	0.18	0.17	0.00	0.00
		0.17	0.19	0.12	0.29	0.30	0.24	0.17	0.20	0.20	0.00	0.19
		0.21	0.23	0.14	0.31	0.34	0.29	0.21	0.23	0.23	0.00	0.34
Resources & groundwater protection	w.2	0.20	0.19	0.21	0.21	0.06	0.13	0.21	0.29	0.12	0.40	0.06
		0.24	0.22	0.24	0.23	0.10	0.17	0.25	0.31	0.15	0.43	0.24
		0.29	0.26	0.28	0.26	0.14	0.21	0.29	0.34	0.18	0.48	0.48
Water supply	w.3	0.31	0.26	0.31	0.33	0.29	0.30	0.25	0.34	0.23	0.35	0.23
		0.34	0.28	0.34	0.36	0.33	0.34	0.28	0.37	0.24	0.39	0.33
		0.38	0.30	0.37	0.38	0.37	0.38	0.31	0.39	0.27	0.43	0.43
Social acceptance	w.4	0.00	0.11	0.06	0.02	0.03	0.09	0.08	0.02	0.17	0.00	0.00
		0.00	0.13	0.08	0.04	0.05	0.12	0.11	0.03	0.20	0.00	0.08
		0.00	0.15	0.11	0.06	0.07	0.15	0.15	0.04	0.23	0.00	0.23
Costs	w.5	0.20	0.16	0.18	0.07	0.19	0.10	0.16	0.07	0.20	0.13	0.07
		0.24	0.18	0.22	0.09	0.23	0.14	0.19	0.09	0.22	0.17	0.18
		0.29	0.21	0.26	0.11	0.28	0.18	0.24	0.11	0.23	0.22	0.29
Rehabilita-tion burden	w.1.1	0.59	0.63	0.23	0.54	0.53	0.56	0.41	0.77	0.63	0.00	0.00
		0.63	0.67	0.26	0.56	0.56	0.59	0.44	0.80	0.67	0.00	0.52
		0.67	0.71	0.29	0.57	0.59	0.63	0.47	0.83	0.71	0.00	0.83
Flexibility	w.1.2	0.33	0.29	0.71	0.43	0.41	0.38	0.53	0.17	0.29	0.00	0.00
		0.38	0.33	0.74	0.44	0.44	0.41	0.56	0.20	0.33	0.00	0.38
		0.41	0.38	0.77	0.46	0.47	0.44	0.59	0.23	0.38	0.00	0.77
Groundwater protection	w.2.1	0.63	0.67	0.71	0.71	0.38	0.63	0.77	0.63	0.80	1.00	0.38
		0.67	0.71	0.74	0.74	0.41	0.67	0.80	0.67	0.83	1.00	0.73
		0.71	0.77	0.77	0.77	0.44	0.71	0.83	0.71	0.87	1.00	1.00
Energy consumption	w.2.2	0.29	0.23	0.23	0.23	0.56	0.29	0.17	0.29	0.13	0.00	0.00
		0.33	0.29	0.26	0.26	0.59	0.33	0.20	0.33	0.17	0.00	0.28
		0.38	0.33	0.29	0.29	0.63	0.38	0.23	0.38	0.20	0.00	0.63
Dw supply	w.3.1	0.36	0.40	0.48	0.48	0.28	0.67	0.33	0.59	0.42	0.36	0.28
		0.38	0.43	0.51	0.50	0.32	0.74	0.36	0.67	0.45	0.37	0.48
		0.42	0.45	0.54	0.53	0.36	0.83	0.38	0.77	0.50	0.38	0.83
Hw supply	w.3.2	0.27	0.29	0.26	0.08	0.37	0.07	0.29	0.23	0.32	0.36	0.07
		0.31	0.32	0.29	0.10	0.40	0.11	0.32	0.33	0.36	0.37	0.29
		0.35	0.35	0.33	0.13	0.43	0.15	0.36	0.41	0.41	0.38	0.43
Ffw supply	w.3.3	0.27	0.22	0.18	0.38	0.24	0.08	0.29	0.00	0.14	0.23	0.00
		0.31	0.26	0.20	0.40	0.28	0.15	0.32	0.00	0.18	0.26	0.24
		0.35	0.29	0.23	0.43	0.32	0.21	0.36	0.00	0.23	0.29	0.43
Dw quantity	w.3.1.1	0.17	0.25	0.18	0.18	0.26	0.21	0.24	0.00	0.21	0.17	0.00
		0.22	0.27	0.20	0.22	0.31	0.23	0.27	0.04	0.25	0.21	0.22
		0.26	0.29	0.23	0.26	0.36	0.25	0.31	0.08	0.29	0.25	0.36
Dw reliability	w.3.1.2	0.40	0.29	0.26	0.28	0.29	0.33	0.31	0.15	0.29	0.38	0.15
		0.43	0.31	0.29	0.32	0.33	0.35	0.33	0.22	0.33	0.42	0.33
		0.48	0.33	0.33	0.36	0.37	0.38	0.36	0.29	0.38	0.45	0.48
Dw quality	w.3.1.3	0.30	0.40	0.48	0.43	0.33	0.40	0.37	0.67	0.38	0.33	0.30
		0.35	0.42	0.51	0.46	0.36	0.42	0.39	0.74	0.42	0.38	0.45
		0.39	0.43	0.54	0.50	0.40	0.43	0.42	0.83	0.45	0.42	0.83
Hw quantity	w.3.2.1	0.27	0.27	0.18	0.23	0.38	0.16	0.25	0.14	0.21	0.29	0.14
		0.32	0.30	0.20	0.26	0.43	0.21	0.29	0.20	0.25	0.32	0.28
		0.36	0.32	0.23	0.29	0.48	0.26	0.33	0.26	0.29	0.36	0.48
Hw reliability	w.3.2.2	0.42	0.43	0.26	0.36	0.43	0.48	0.38	0.42	0.29	0.33	0.26
		0.45	0.45	0.29	0.37	0.48	0.53	0.42	0.45	0.33	0.36	0.42
		0.50	0.48	0.33	0.38	0.53	0.59	0.45	0.50	0.38	0.38	0.59

Objective	Weight no.	SH1 min mean max	SH2 min mean max	SH3 min mean max	SH4 min mean max	SH5 min mean max	SH6 min mean max	SH7 min mean max	SH8 min mean max	SH9 min mean max	SH10 min mean max	over-all min mean max
Hw quality	w.3.2.3	0.18 0.23 0.27	0.23 0.25 0.27	0.48 0.51 0.54	0.36 0.37 0.38	0.05 0.10 0.14	0.21 0.26 0.32	0.25 0.29 0.33	0.30 0.34 0.38	0.38 0.42 0.45	0.29 0.32 0.36	0.05 0.31 0.54
Dw esthetic quality	w.3.1.3.1	0.27 0.32 0.36	0.33 0.35 0.38	0.19 0.21 0.23	0.29 0.31 0.33	0.33 0.33 0.33	0.33 0.33 0.33	0.22 0.27 0.32	0.07 0.10 0.13	0.35 0.40 0.45	0.38 0.40 0.42	0.07 0.30 0.45
Dw microbial & hygienic quality	w.3.1.3.2	0.42 0.45 0.50	0.40 0.42 0.43	0.44 0.47 0.49	0.33 0.34 0.36	0.33 0.33 0.33	0.33 0.33 0.33	0.40 0.45 0.50	0.63 0.67 0.71	0.45 0.50 0.56	0.38 0.40 0.42	0.33 0.44 0.71
Dw physico-chemical quality	w.3.1.3.3	0.18 0.23 0.27	0.21 0.23 0.25	0.30 0.33 0.35	0.33 0.34 0.36	0.33 0.33 0.33	0.33 0.33 0.33	0.23 0.27 0.35	0.20 0.23 0.27	0.05 0.10 0.15	0.17 0.20 0.23	0.05 0.26 0.36
Hw esthetic quality	w.3.2.3.1	0.24 0.29 0.33	0.45 0.48 0.50	0.19 0.21 0.23	0.24 0.28 0.32	0.63 0.71 0.83	0.38 0.42 0.45	0.33 0.38 0.41	0.33 0.43 0.44	0.33 0.38 0.41	0.50 0.53 0.56	0.19 0.41 0.83
Hw microbial & hygienic quality	w.3.2.3.2	0.43 0.48 0.53	0.29 0.31 0.33	0.44 0.47 0.49	0.37 0.40 0.43	0.07 0.14 0.21	0.25 0.29 0.33	0.59 0.63 0.67	0.56 0.57 0.67	0.59 0.63 0.67	0.44 0.47 0.50	0.07 0.44 0.67
Hw physico-chemical quality	w.3.2.3.3	0.19 0.24 0.29	0.19 0.21 0.24	0.30 0.33 0.35	0.28 0.32 0.36	0.07 0.14 0.21	0.25 0.29 0.33	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.15 0.36
Dw hygiene	w.3.1.3.2.1	0.63 0.67 0.71	0.50 0.52 0.54	0.50 0.50 0.50	0.50 0.50 0.50	0.53 0.56 0.59	0.77 0.83 0.91	0.63 0.65 0.67	0.77 0.83 0.91	0.59 0.67 0.77	1.00 1.00 1.00	0.50 0.68 1.00
Dw microbial regrowth	w.3.1.3.2.2	0.29 0.33 0.38	0.46 0.48 0.50	0.50 0.50 0.50	0.50 0.50 0.50	0.41 0.44 0.47	0.09 0.17 0.23	0.33 0.35 0.38	0.09 0.17 0.23	0.23 0.33 0.41	0.00 0.00 0.00	0.00 0.33 0.50
Hw hygiene	w.3.2.3.2.1	0.67 0.71 0.77	0.53 0.56 0.59	0.51 0.53 0.56	0.50 0.50 0.50	0.50 0.50 0.50	0.67 0.71 0.77	0.77 0.77 0.77	0.77 0.83 0.91	0.59 0.67 0.77	1.00 1.00 1.00	0.50 0.68 1.00
Hw microbial regrowth	w.3.2.3.2.2	0.23 0.29 0.33	0.41 0.44 0.47	0.44 0.47 0.49	0.50 0.50 0.50	0.50 0.50 0.50	0.23 0.29 0.33	0.23 0.23 0.23	0.09 0.17 0.23	0.23 0.33 0.41	0.00 0.00 0.00	0.00 0.32 0.50
Dw inorganics	w.3.1.3.3.1	0.33 0.33 0.33	0.20 0.22 0.25	0.35 0.36 0.37	0.33 0.35 0.37	0.33 0.33 0.33	0.00 0.33 1.00	0.17 0.18 0.20	0.40 0.45 0.53	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.27 1.00
Dw pesticides	w.3.1.3.3.2	0.33 0.33 0.33	0.38 0.41 0.43	0.35 0.36 0.37	0.33 0.35 0.37	0.33 0.33 0.33	0.00 0.33 1.00	0.40 0.41 0.42	0.17 0.23 0.29	0.00 0.00 0.00	0.50 0.53 0.56	0.00 0.34 1.00
Dw micropollutants	w.3.1.3.3.3	0.33 0.33 0.33	0.33 0.37 0.40	0.26 0.28 0.30	0.26 0.30 0.33	0.33 0.33 0.33	0.00 0.33 1.00	0.40 0.41 0.42	0.24 0.32 0.39	1.00 1.00 1.00	0.44 0.47 0.50	0.00 0.42 1.00
Hw inorganics	w.3.2.3.3.1	0.33 0.33 0.33	0.12 0.15 0.19	0.36 0.37 0.39	0.33 0.35 0.37	0.33 0.33 0.33	0.00 0.33 1.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.20 1.00
Hw pesticides	w.3.2.3.3.2	0.33 0.33 0.33	0.56 0.61 0.67	0.36 0.37 0.39	0.33 0.35 0.37	0.33 0.33 0.33	0.00 0.33 1.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.24 1.00
Hw micropollutants	w.3.2.3.3.3	0.33 0.33 0.33	0.19 0.24 0.29	0.23 0.26 0.27	0.26 0.30 0.33	0.33 0.33 0.33	0.00 0.33 1.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.00 0.00	0.00 0.19 1.00
Ffw reliability	w.3.3.1	0.56 0.59 0.63	0.63 0.67 0.71	0.56 0.57 0.59	0.50 0.54 0.59	0.56 0.59 0.63	0.63 0.67 0.71	0.53 0.54 0.56	0.00 0.00 0.00	0.59 0.67 0.77	0.67 0.74 0.83	0.00 0.56 0.83
Ffw quantity	w.3.3.2	0.38 0.41 0.44	0.29 0.33 0.38	0.41 0.43 0.44	0.41 0.46 0.50	0.38 0.41 0.44	0.29 0.33 0.38	0.44 0.46 0.47	0.00 0.00 0.00	0.23 0.33 0.41	0.17 0.26 0.33	0.00 0.34 0.50
Operational & management	w.4.1	0.00 0.00 0.00	0.23 0.25 0.28	0.20 0.22 0.24	0.06 0.07 0.09	0.21 0.24 0.28	0.71 0.77 0.83	0.30 0.34 0.38	0.31 0.35 0.40	0.24 0.28 0.33	0.00 0.00 0.00	0.00 0.25 0.83

Objective	Weight no.	SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	over-all
		min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max	min mean max
quality												
Co-determi-nation	w.4.2	0.00 0.00 0.00	0.14 0.16 0.19	0.16 0.18 0.21	0.01 0.02 0.03	0.18 0.21 0.25	0.17 0.23 0.29	0.13 0.17 0.21	0.06 0.09 0.12	0.03 0.06 0.09	0.00 0.00 0.00	0.00 0.11 0.29
Resources autonomy	w.4.3	0.00 0.00 0.00	0.17 0.19 0.22	0.13 0.15 0.18	0.28 0.30 0.32	0.00 0.00 0.00	0.00 0.00 0.00	0.06 0.08 0.11	0.23 0.28 0.33	0.03 0.06 0.09	0.00 0.00 0.00	0.00 0.11 0.33
Time demand	w.4.4	0.00 0.00 0.00	0.10 0.11 0.14	0.09 0.12 0.14	0.23 0.25 0.28	0.21 0.24 0.28	0.00 0.00 0.00	0.06 0.08 0.11	0.00 0.00 0.00	0.18 0.22 0.28	0.00 0.00 0.00	0.00 0.10 0.28
Areal demand	w.4.5	0.00 0.00 0.00	0.05 0.08 0.11	0.13 0.15 0.18	0.17 0.19 0.22	0.00 0.00 0.00	0.00 0.00 0.00	0.06 0.08 0.11	0.10 0.14 0.19	0.18 0.22 0.28	0.00 0.00 0.00	0.00 0.09 0.28
Unnecessary disturbance from road works	w.4.6	0.00 0.00 0.00	0.17 0.20 0.24	0.16 0.18 0.21	0.14 0.16 0.18	0.28 0.30 0.33	0.00 0.00 0.00	0.19 0.24 0.29	0.10 0.14 0.19	0.11 0.17 0.24	0.00 0.00 0.00	0.00 0.14 0.33
Annual costs	w.5.1	0.53 0.56 0.59	0.33 0.38 0.41	0.50 0.50 0.50	0.41 0.44 0.47	0.33 0.38 0.41	0.63 0.67 0.71	0.63 0.67 0.71	0.50 0.50 0.50	1.00 1.00 1.00	0.23 0.31 0.38	0.23 0.54 1.00
Cost increase	w.5.2	0.41 0.44 0.47	0.59 0.63 0.67	0.50 0.50 0.50	0.53 0.56 0.59	0.59 0.63 0.67	0.29 0.33 0.38	0.29 0.33 0.38	0.50 0.50 0.50	0.00 0.00 0.00	0.63 0.69 0.77	0.29 0.33 0.38

Table SI4.2: Ranking of objectives and relevance from the online survey for ten stakeholders (see Tab. SI1.1). "1" indicates first rank = most important objective, and decreasing ranks indicate objectives of decreasing importance. Irrelevant objectives that could be dismissed according to the respective stakeholder are grey shaded (rank 0). The objectives were ranked top-down following the hierarchical structure of the objectives hierarchy. The sub-objectives of microbial & hygienic quality and of physico-chemical quality of drinking water and household water were not ranked in the online survey. SH = stakeholder, dw = drinking water, hw= household water, ffw = firefighting water, ∑w=0 sum of irrelevant objectives.

		Stakeholder										Overall		
Objective	Weight	SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	min	mean	max
Intergenerational equity	w.1	4	3	5	3	1	2	3	3	3	4	1	3.1	5
Resources & groundwater protection	w.2	2	2	2	2	4	3	2	2	4	1	1	2.4	4
Water supply	w.3	1	1	1	1	1	1	1	1	1	2	1	1.1	2
Social acceptance	w.4	5	5	4	4	4	4	5	5	2	5	2	4.1	5
Costs	w.5	2	3	3	4	3	4	4	4	5	3	2	3.5	5
Rehabilitation burden	w.1.1	**0**	**0**	**0**	1	1	1	2	1	**0**	**0**	0	0.6	2
Flexibility	w.1.2	**0**	**0**	**0**	2	2	2	1	2	**0**	**0**	0	0.9	2
Groundwater protection	w.2.1	1	1	1	1	2	2	1	1	1	1	1	1.2	2
Energy consump-tion	w.2.2	2	2	2	2	1	1	2	2	2	2	1	1.8	2
Dw supply	w.3.1	1	1	1	1	2	1	1	1	1	1	1	1.1	2
Hw supply	w.3.2	2	1	2	3	1	3	1	2	2	1	1	1.8	3
Ffw supply	w.3.3	2	3	3	2	3	2	3	3	3	1	1	2.5	3

Objective	Weight	Stakeholder SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	Overall min	mean	max
Dw quantity	w.3.1.1	3	3	3	3	3	3	3	3	3	3	3	3.0	3
Dw reliability	w.3.1.2	1	1	2	2	2	2	2	2	2	1	1	1.7	2
Dw quality	w.3.1.3	2	1	1	1	1	1	1	1	1	1	1	1.1	2
Hw quantity	w.3.2.1	3	2	3	3	2	3	3	3	3	3	2	2.9	3
Hw reliability	w.3.2.2	1	1	2	2	1	2	2	1	2	1	1	1.5	2
Hw quality	w.3.2.3	2	1	1	1	3	1	1	2	1	1	1	1.4	3
Dw esthetic quality	w.3.1.3.1	**0**	1	3	3	1	1	2	3	2	1	0	1.7	3
Dw microbial & hygienic quality	w.3.1.3.2	**0**	1	1	1	1	1	1	1	1	1	0	0.9	1
Dw physico-chemical quality	w.3.1.3.3	**0**	3	2	2	1	1	2	2	3	1	0	1.7	3
Hw esthetic quality	w.3.2.3.1	**0**	1	3	3	1	1	2	2	2	1	0	1.6	3
Hw microbial & hygienic quality	w.3.2.3.2	**0**	1	1	1	1	1	1	1	1	1	0	0.9	1
Hw physico-chemical quality	w.3.2.3.3	**0**	3	2	2	1	1	2	3	3	1	0	1.8	3
Ffw reliability	w.3.3.1	1	1	1	1	1	**0**	1	**0**	1	1	0	0.8	1
Ffw quantity	w.3.3.2	2	2	2	2	2	**0**	1	**0**	2	1	0	1.4	2
Operational & management quality	w.4.1	**0**	**0**	**0**	**0**	4	1	**0**	**0**	2	**0**	0	0.7	4
Co-determination	w.4.2	**0**	**0**	**0**	**0**	3	2	**0**	**0**	4	**0**	0	0.9	4
Resources autonomy	w.4.3	**0**	**0**	**0**	**0**	6	4	**0**	**0**	5	**0**	0	1.5	6
Time demand	w.4.4	**0**	**0**	**0**	**0**	2	2	**0**	**0**	3	**0**	0	0.7	3
Areal demand	w.4.5	**0**	**0**	**0**	**0**	5	4	**0**	**0**	6	**0**	0	1.5	6
Unnecessary disturbance from road works	w.4.6	**0**	**0**	**0**	**0**	1	4	**0**	**0**	1	**0**	0	0.6	4
Annual costs	w.5.1	1	2	1	2	2	1	1	1	**0**	2	0	1.3	2
Cost increase	w.5.2	1	1	1	1	1	2	2	2	**0**	1	0	1.2	2
	∑w=0	14	8	8	6	0	2	6	8	4	8			

Table SI4.3: Comparison of ranks and relevance of objectives in face-to-face interviews (Tab. SI4.1) and the online survey (Tab. SI4.2). SH = stakeholder, dw = drinking water, hw= household water, ffw = firefighting water, ∑ = sum relevant objectives over 10 stakeholders (percentage given in parenthesis, the maximum no. of relevant objectives is 340 = 100 %).

Objective	Weight	Survey rank min	Survey rank mean	Survey rank max	Survey relevance # no of SH for which relevant	Interview rank min	Interview rank mean	Interview rank max	Interview relevance # no of SH for which relevant
Intergenerational equity	w.1	1	3.1	5	5	2	3.1	4	9
Resources & groundwater protection	w.2	1	2.4	4	10	1	2.6	5	10
Water supply	w.3	1	1.1	2	10	1	1.1	2	10
Social acceptance	w.4	2	4.3	5	3	2	4.7	5	8
Costs	w.5	2	3.5	5	9	2	3.2	4	10
Rehabilitation burden	w.1.1	1	1.1	2	5	1	1.2	2	9
Flexibility	w.1.2	1	1.4	2	4	1	1.7	2	9
Groundwater protection	w.2.1	1	1.2	2	8	1	1.1	2	10
Energy consumption	w.2.2	1	1.8	2	4	1	1.9	2	9
Dw supply	w.3.1	1	1.1	2	10	1	1.1	2	10
Hw supply	w.3.2	1	1.8	3	7	1	2.0	3	10
Ffw supply	w.3.3	1	2.5	3	8	2	2.6	3	9
Dw quantity	w.3.1.1	3	3	3	9	3	3.0	3	10
Dw reliability	w.3.1.2	1	1.7	2	9	1	1.8	2	10
Dw quality	w.3.1.3	1	1.1	2	9	1	1.2	2	10
Hw quantity	w.3.2.1	2	2.9	3	5	2	2.5	3	10
Hw reliability	w.3.2.2	1	1.5	2	6	1	1.2	2	10
Hw quality	w.3.2.3	1	1.4	3	6	1	2.0	3	10
Dw esthetic quality	w.3.1.3.1	1	1.8	3	9	1	2.1	3	10
Dw microbial & hygienic quality	w.3.1.3.2	1	1	1	9	1	1.0	1	10
Dw physico-chemical quality	w.3.1.3.3	1	1.8	3	9	1	2.1	3	10
Hw esthetic quality	w.3.2.3.1	1	1.7	3	5	1	1.8	3	10
Hw microbial & hygienic quality	w.3.2.3.2	1	1	1	6	1	1.4	2	10
Hw physico-chemical quality	w.3.2.3.3	1	1.9	3	4	2	2.6	3	6
Ffw reliability	w.3.3.1	1	1	1	8	1	1.0	1	9
Ffw quantity	w.3.3.2	1	1.6	2	6	1	1.9	2	9
Operational & management quality	w.4.1	1	1.4	4	1	1	1.5	5	8
Co-determination	w.4.2	1	1.6	4	2	1	3.3	6	8
Resources autonomy	w.4.3	1	2.2	6	1	1	2.9	5	6
Time demand	w.4.4	1	1.4	3	1	1	3.2	6	6
Areal demand	w.4.5	1	2.2	6	0	1	3.2	6	6
Unnecessary disturbance from road works	w.4.6	1	1.3	4	2	1	2.3	4	7
Annual costs	w.5.1	0	1.3	2	7	1	1.4	2	10
Cost increase	w.5.2	0	1.2	2	8	1	1.4	2	9
				∑	205 (60.3%)			∑	307 (90.3%)

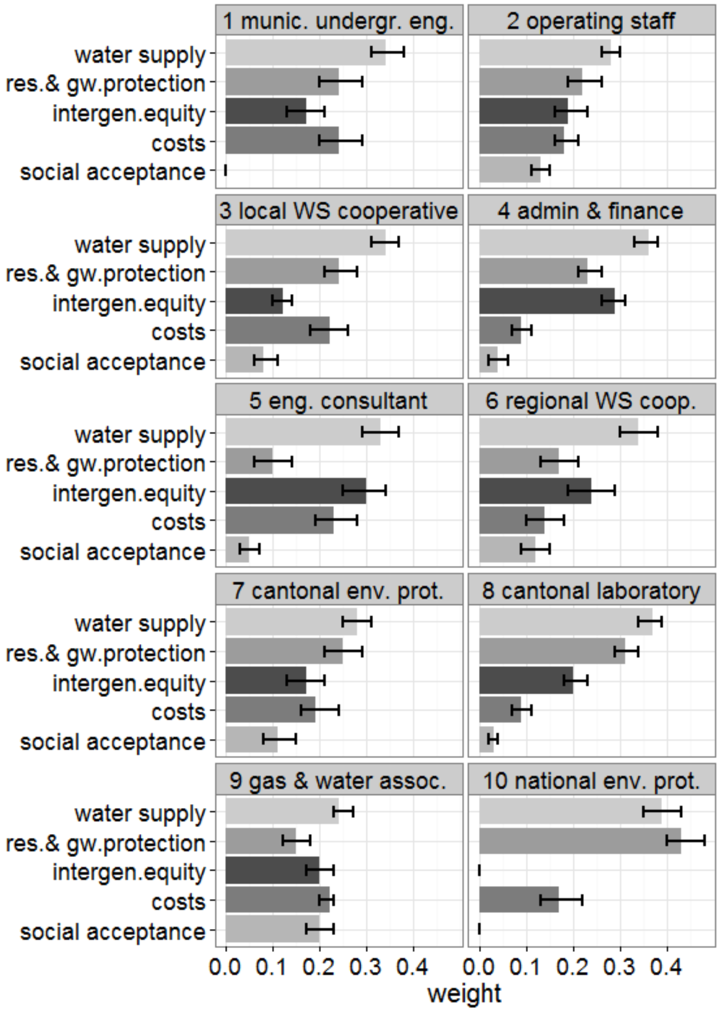


Figure SI4.1: Weights of the sub-objectives of 'good water supply infrastructure', i.e. the five highest-level objectives for the ten stakeholders (see Tab. SI1.1)

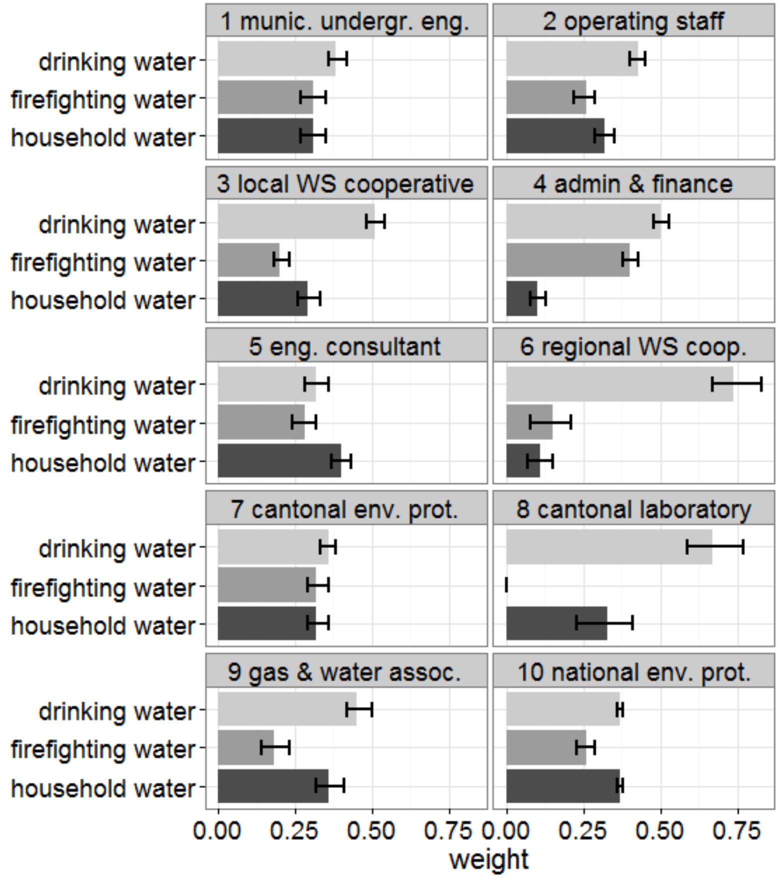


Figure SI4.2: Weights of the sub-objectives of 'good water supply'

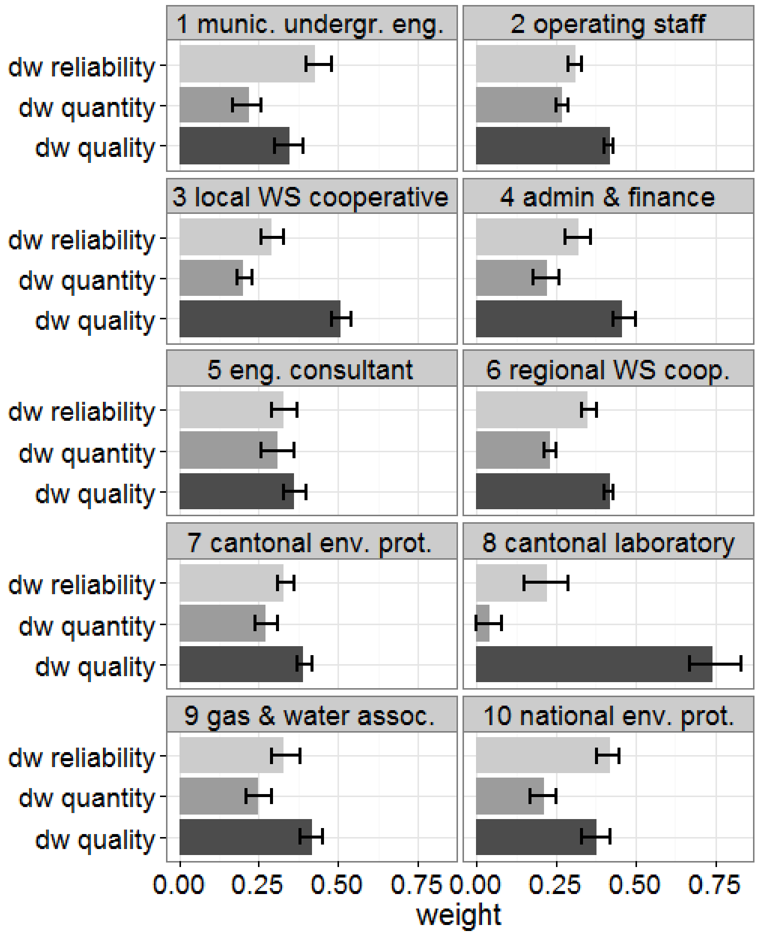


Figure SI4.3: Weights of the sub-objectives of 'good water supply – drinking water'

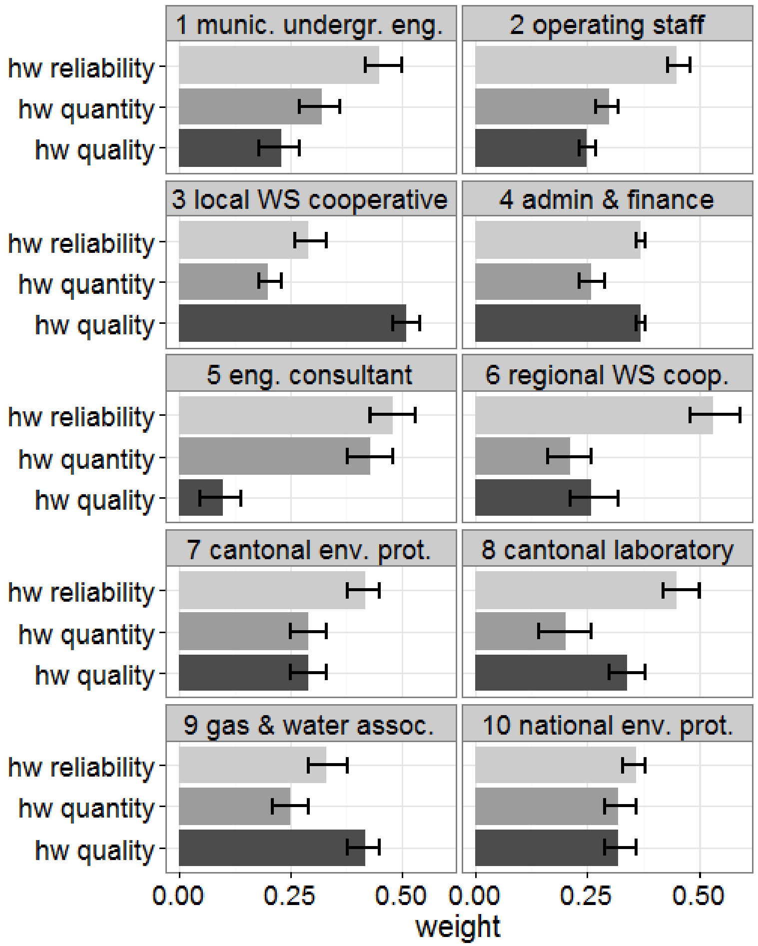


Figure SI4.4: Weights of the sub-objectives of 'good water supply – household water'

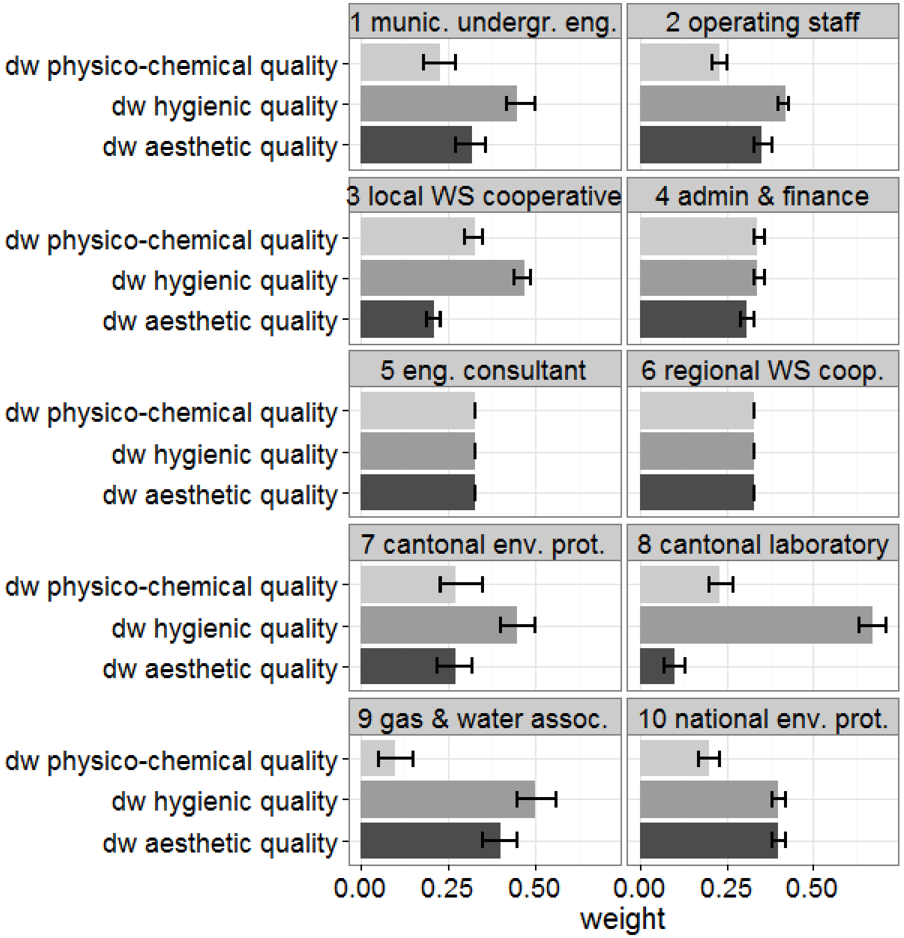


Figure SI4.5: Weights of the sub-objectives of ' good water supply – drinking water-quality'

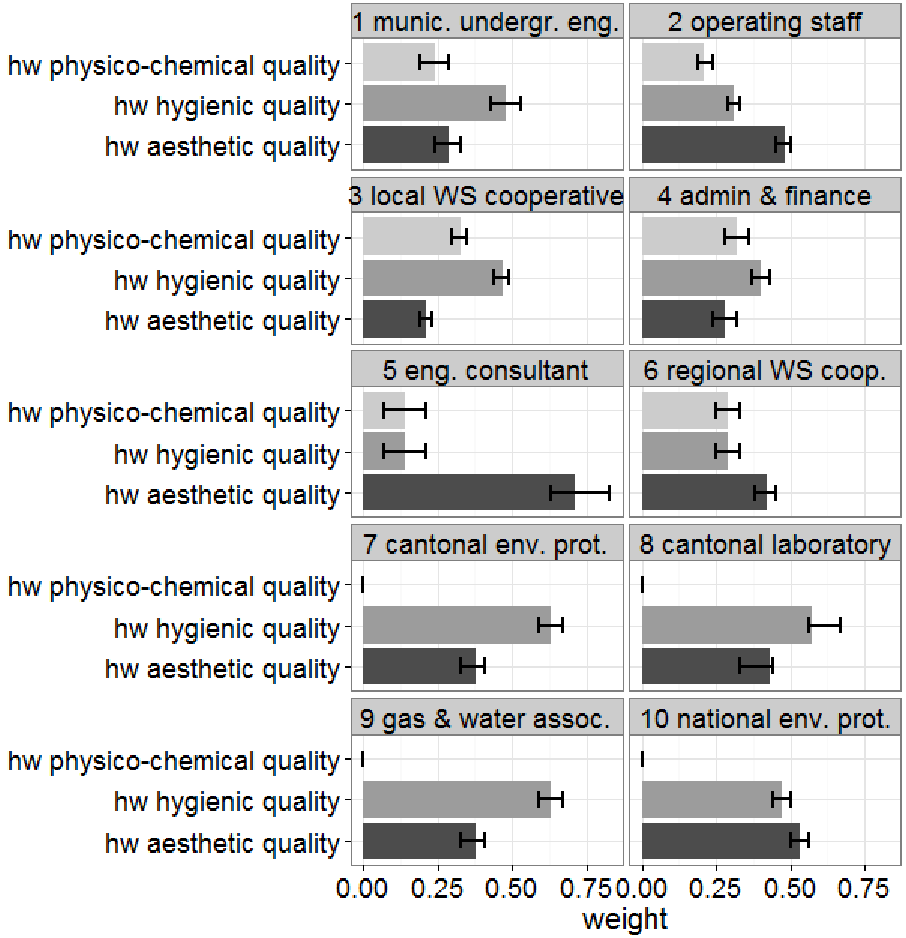


Figure SI4.6: Weights of the sub-objectives of 'good water supply – household water-quality'

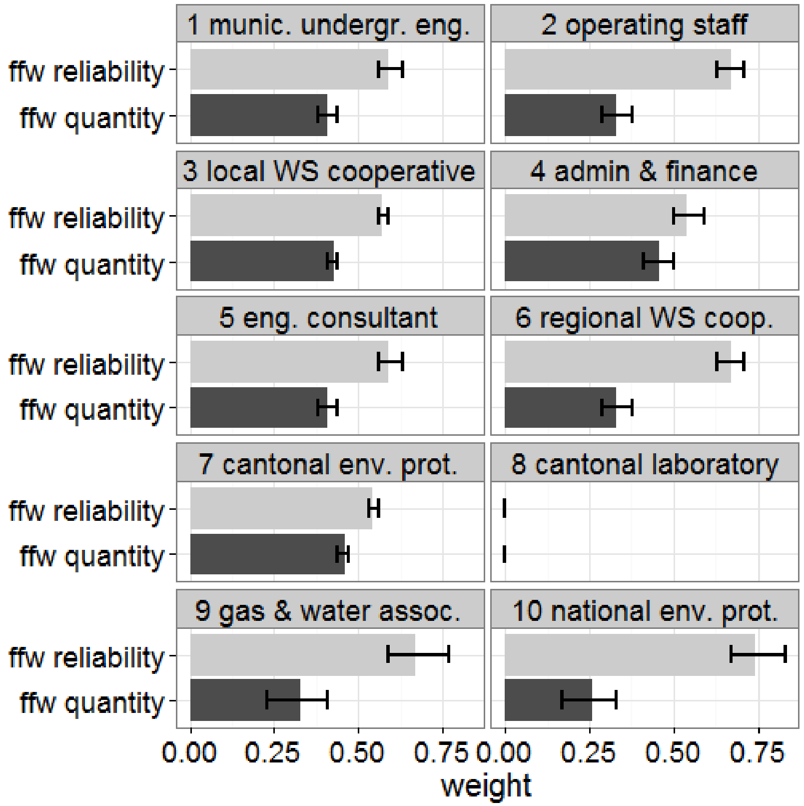


Figure SI4.7: Weights of the sub-objectives of 'good water supply – firefighting water'

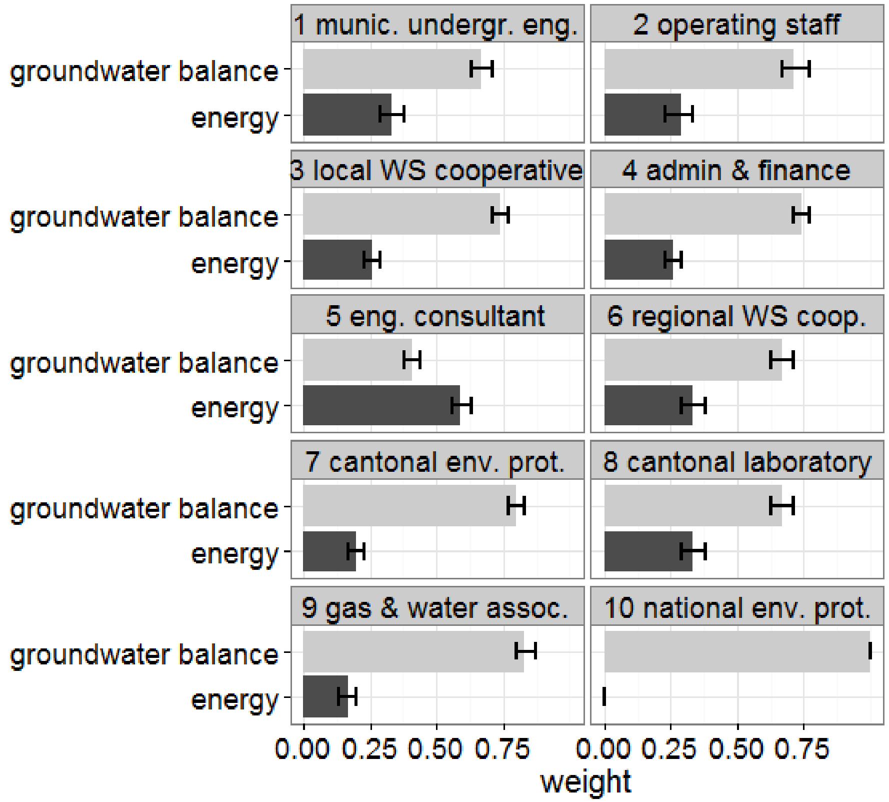


Figure SI4.8: Weights of the sub-objectives of 'good resources and groundwater protection'

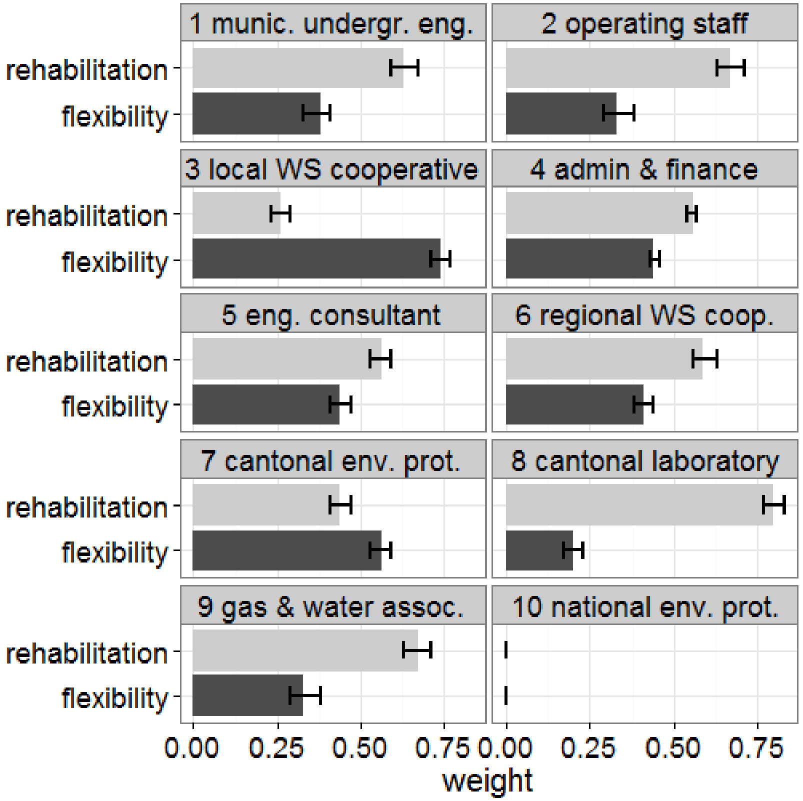


Figure SI4.9: Weights of the sub-objectives of 'high intergenerational equity'

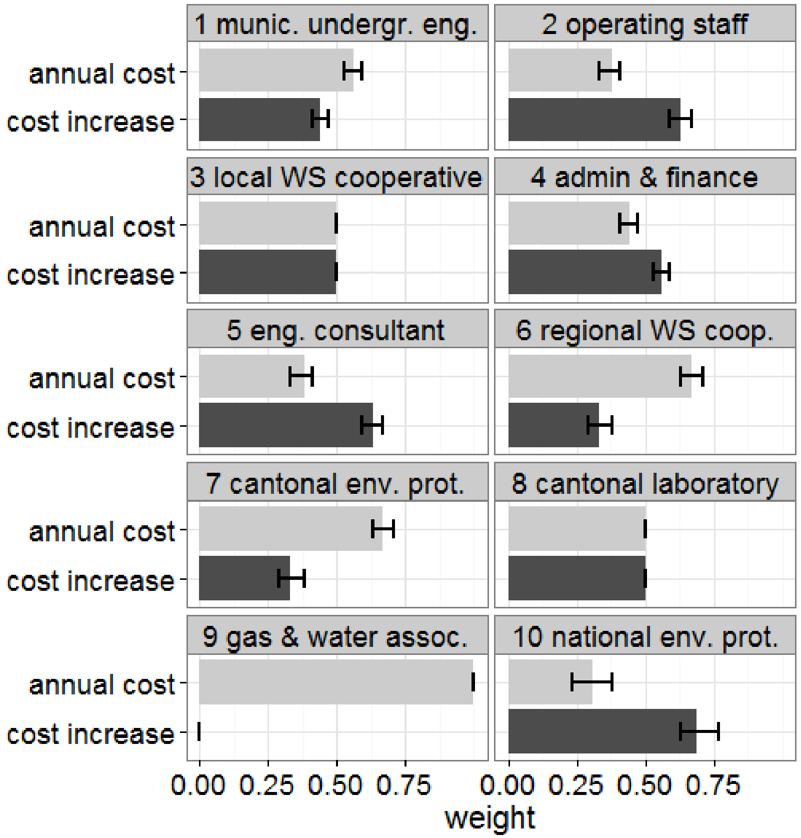


Figure SI4.10: Weights of the sub-objectives of 'low costs'

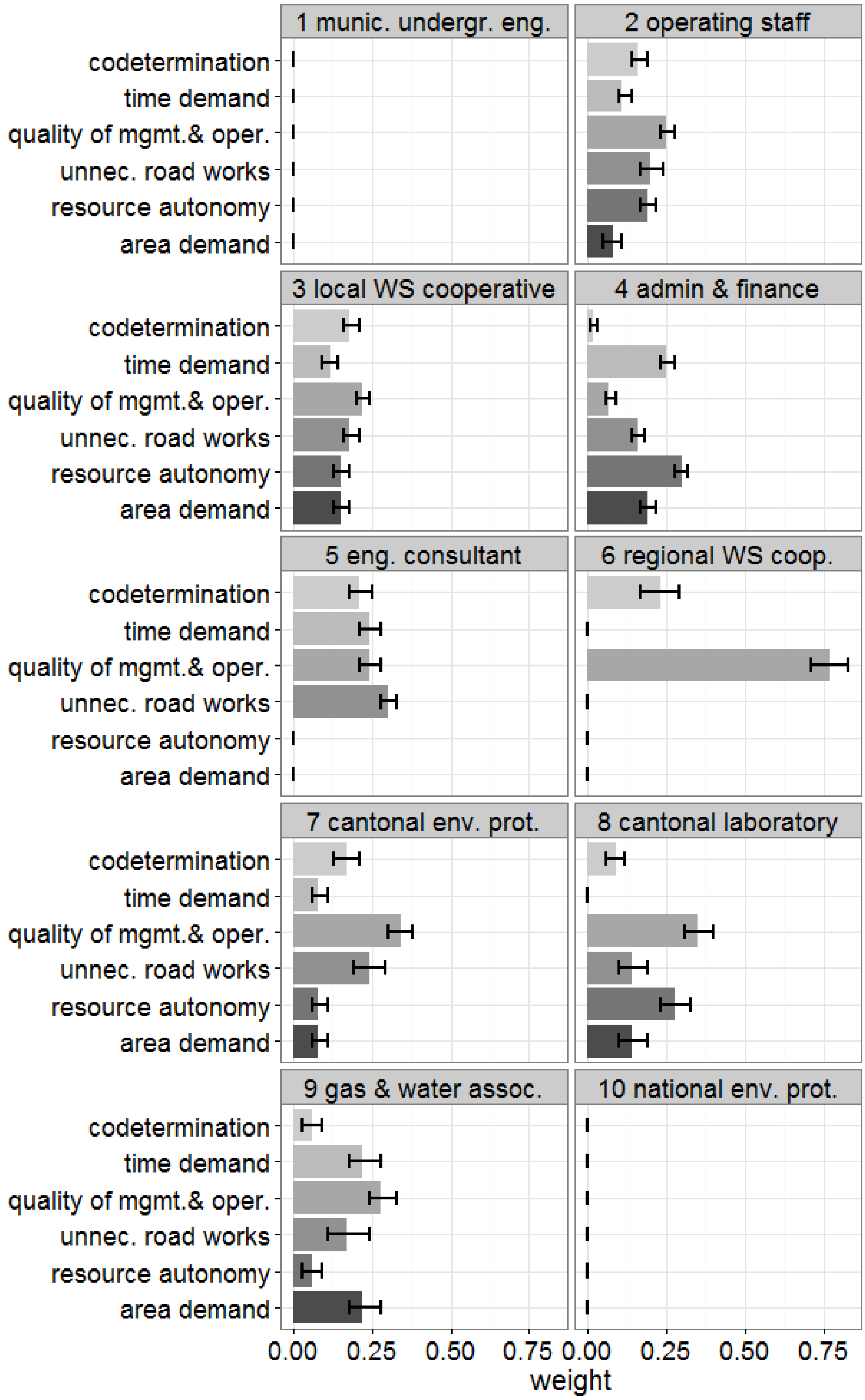


Figure SI4.11: Weights of the sub-objectives of 'high social acceptance'

Marginal value functions

Table SI4.4: Fitted marginal value function curvature parameters (elicited $v_{0.25}$, $v_{0.5}$, $v_{0.75}$ values, and standard deviation of the fit not shown) for ten stakeholders. The bold numbers indicate parameter c_i, which determines the curvature of the function (see Material and methods, main text). These numbers were derived from fitting an exponential function to the elicited "best guess" and range for the 0.25, 0.5, and 0.75 values from the interview partner. For reasons of time, in most cases only a rough indication of the shape of the curvature was elicited, where: c<0…convex, c≈0 …linear (cutoff at ± 0.4), and c>0…concave. "Overall" indicates the number of stakeholders assigned to one of three general shapes of the value function. "Summary" in the last row indicates for each stakeholder how many times one of the three respective shapes was observed.

Attribute	SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	Overall		
											c>0	c≈0	c<0
rehab	-	**0.97**	-	**-0.14**	**1.10**	c>0	c>0	**0.83**	**2.25**	-	6	1	0
adapt	-	-	-	-	-	c>0	c>0	c≈0	c>0	-	3	1	0
gwhh	**1.61**	**0.50**	**0.92**	**0.92**	-	c>0	**0.33**	**0.99**	c>0	**other funct.**	9	0	0
econs	-	-	-	-	c≈0	c≈0	-	c≈0	-	-	0	3	0
vol_dw	-	-	c<0	c>0	-	c<0	c>0	c<0	c<0	c>0	3	0	4
reliab_dw	c>0	c<0	c≈0	c≈0	c>0	c<0	c<0	c<0	c>0	**1.59**	4	2	4
vol_hw	-	-	c<0	c>0	-	-	c>0	c<0	c<0	c>0	3	0	3
reliab_hw	c>0	c<0	c≈0	c≈0	c>0	-	c<0	c<0	c>0	c>0	4	2	3
aes_dw	c<0	c<0	c<0	c<0	-	c<0	c>0	c<0	c<0	c<0	1	0	8
aes_hw	c<0	c<0	c<0	c<0	-	-	c>0	c<0	c<0	c<0	1	0	7
faecal_dw	c<0	c<0	c<0	**-1.31**	**-3.06**	**-4.16**	c<0	**-7.60**	c<0	c<0	0	0	10
cells_dw	c>0	c<0	c>0	c>0	-	-	c>0	c>0	c>0	-	6	0	1
faecal_hw	c<0	c<0	c<0	c<0	-	c<0	c<0	c<0	c<0	c<0	0	0	9
cells_hw	c>0	c<0	c>0	c>0	-	-	c>0	c>0	c>0	-	6	0	1
no3_dw	c≈0	-	-	c>0	-	-	-	c<0	-	-	1	1	1
pest_dw	c≈0	-	-	c>0	-	-	-	c<0	-	-	1	1	1
bta_dw	c≈0	-	-	c>0	-	-	-	c<0	-	-	1	1	1
no3_hw	c≈0	-	-	c>0	-	-	-	-	-	-	1	1	0
pest_hw	c≈0	-	-	c>0	-	-	-	-	-	-	1	1	0
bta_hw	c≈0	-	-	c>0	-	-	-	-	-	-	1	1	0
reliab_ffw	c>0	c<0	c≈0	c≈0	c>0	-	c>0	-	c>0	c>0	5	2	1
vol_ffw	c≈0	-	c≈0	c>0	-	-	c>0	-	c>0	c>0	4	2	0
efqm	-	-	-	c>0	-	c>0	c>0	-	c<0	-	3	0	1
voice	-	-	-	c≈0	-	c>0	c>0	-	c>0	-	3	1	0
auton	-	-	-	c>0	-	-	c>0	-	c>0	-	3	0	0
time	-	-	-	c>0	-	-	c≈0	-	c>0	-	2	1	0
area	-	-	-	c>0	-	-	c>0	-	c>0	-	3	0	0
collab	-	-	-	c>0	c>0	-	c<0	-	c<0	-	2	0	2
costcap	**1.83**	-	c>0	c≈0	-	c>0	c>0	c<0	c>0	c>0	6	1	1
costchange	c>0	c<0	**-0.06**	**0.89**	c<0	c>0	c>0	c<0	-	**0.65**	5	1	3
SUMMARY	c>0:8 c≈0: 7 c<0: 4	c>0: 2 c≈0: 0 c<0: 10	c>0: 4 c≈0: 5 c<0: 6	c>0: 18 c≈0: 6 c<0: 4	c>0: 5 c≈0: 1 c<0: 2	c>0: 7 c≈0: 1 c<0: 5	c>0: 17 c≈0: 1 c<0: 5	c>0: 4 c≈0: 2 c<0: 13	c>0: 14 c≈0: 0 c<0: 8	c>0: 9 c≈0: 0 c<0: 4	88	23	61

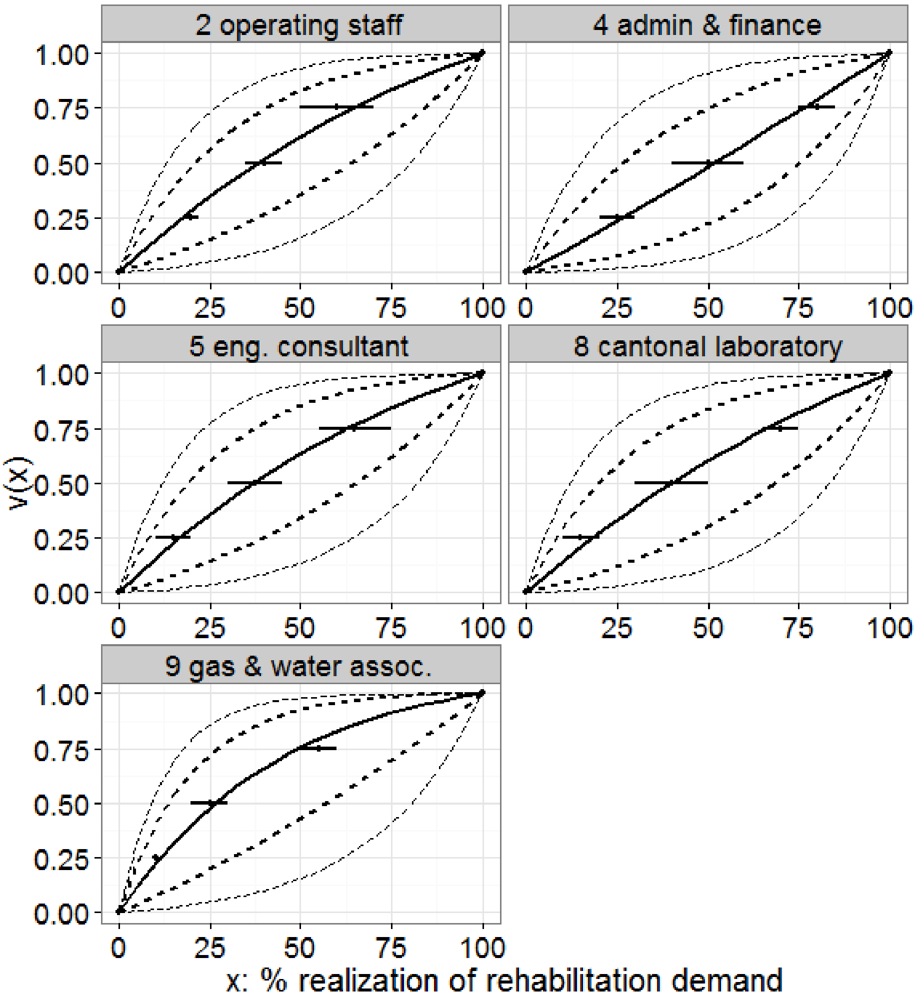


Figure SI4.12a: Elicited value function levels for 'Realization of the rehabilitation demand [%]' (rehab) and fitted distributions for five stakeholders. The value v(x) = 0 on the y-axis indicates that this objective is not at all achieved, and 1 that it is fully achieved. Horizontal intervals show both endpoints and the midpoints (the "best guess") as stated by the decision makers. The solid, black curve represents the value function using the mean exponential parameter μ_c, dashed lines the 95% confidence intervals of μ_c considering half the standard deviation of the fit (used for uncertainty propagation). For comparison, the 95% confidence intervals considering the full standard deviation of the fit is also plotted (thin dashed line). For the remaining stakeholders, only the approximate shape of the curvature was elicited (see Tab. SI4.4).

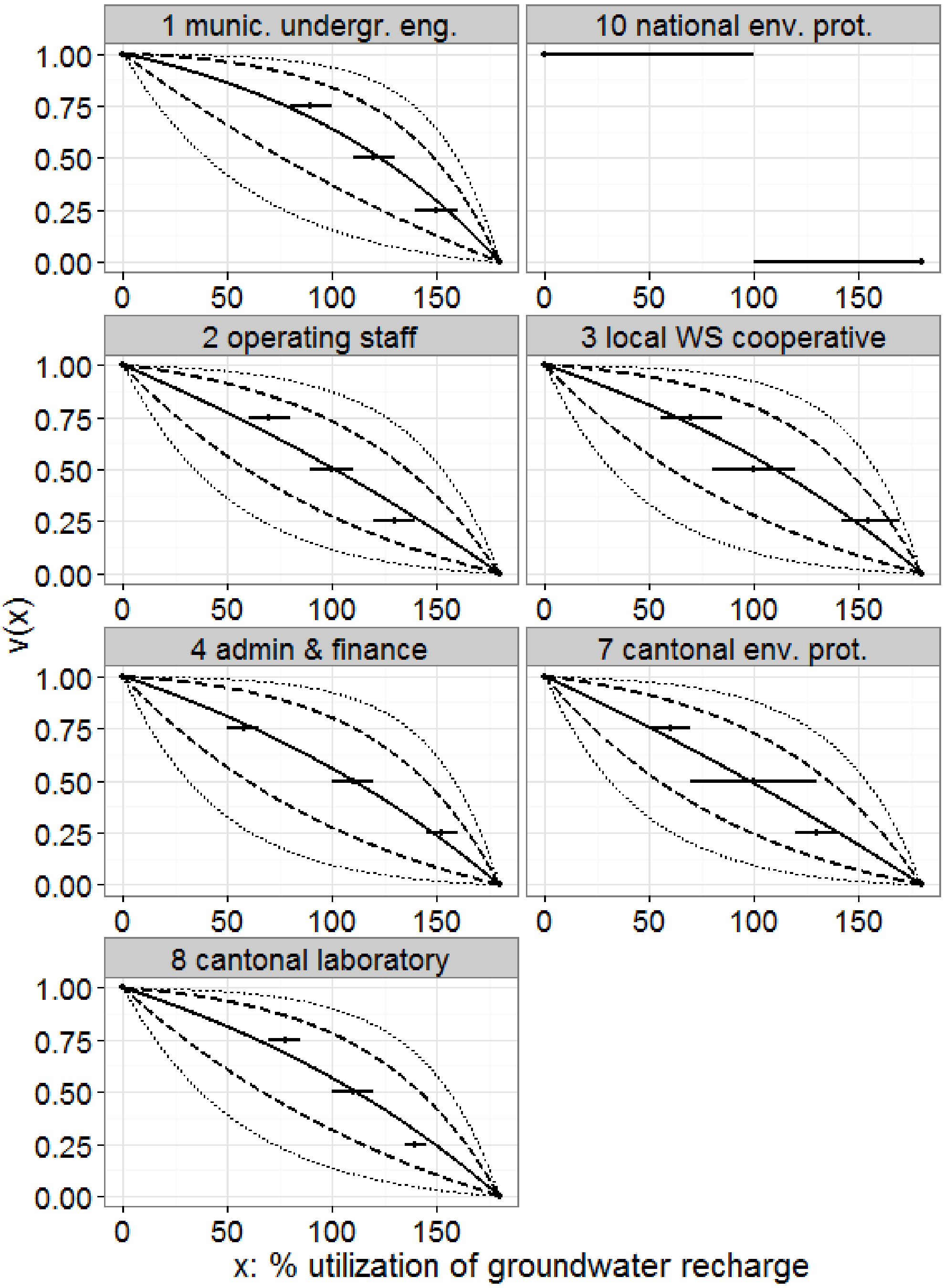


Figure SI4.12b: Elicited value function levels for '% Utilization of groundwater recharge [%]' (gwhh) and fitted distributions for seven stakeholders. For detailed description see Fig. SI4.12a.

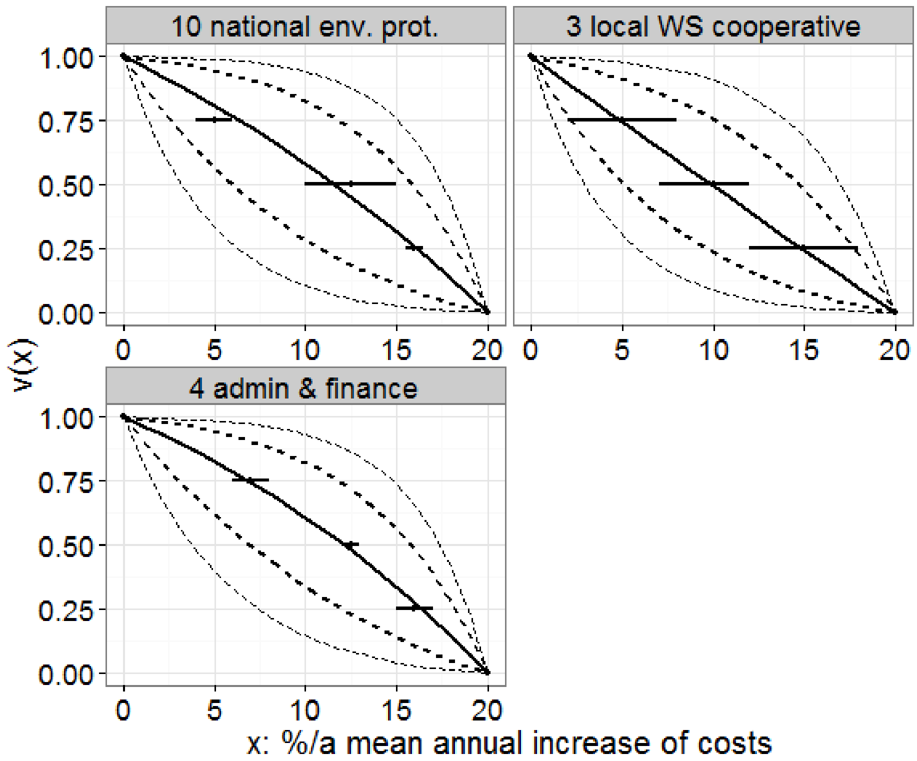


Figure SI4.12c: Elicited value function levels for 'Mean annual (linear) increase of costs [%/a]' (costchange) and fitted distributions for three stakeholders. For detailed description see Fig. SI4.12a.

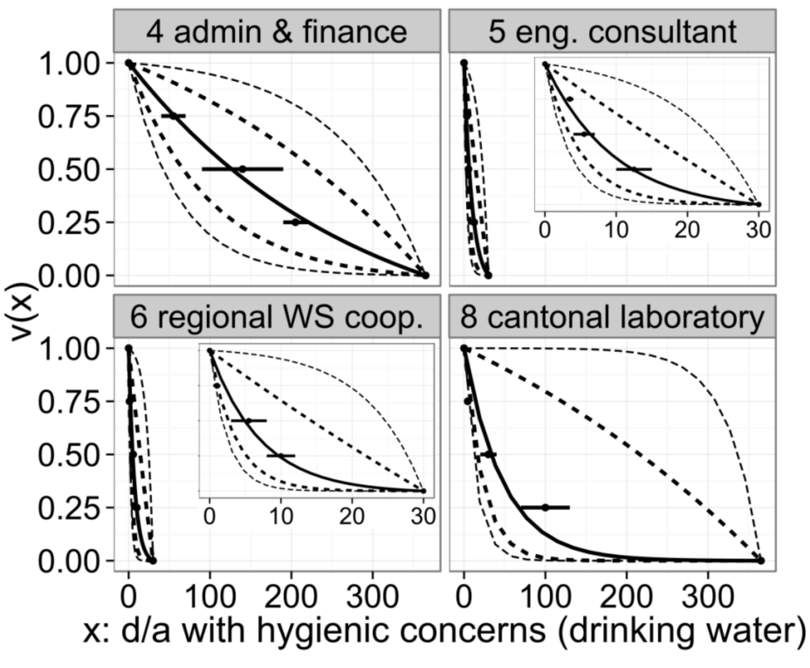


Figure SI4.12d: Elicited value function levels for 'Days per year with hygienic concerns (drinking water)' (faecal_dw) and fitted distributions for four stakeholders. For detailed description see Fig. SI4.12a.

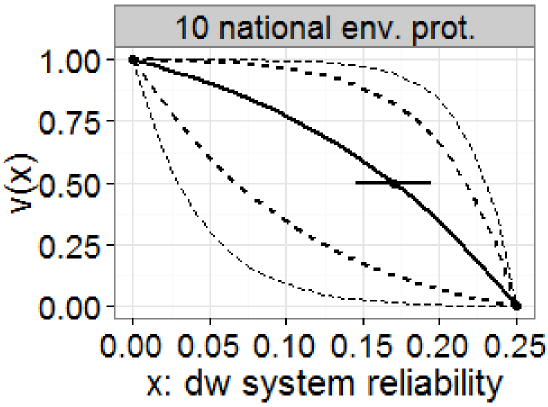


Figure SI4.12e: Elicited value function instances for 'Drinking water system reliability' (reliab_dw) and fitted distributions used for uncertainty propagation for one stakeholder. For detailed description see Fig. SI4.12a.

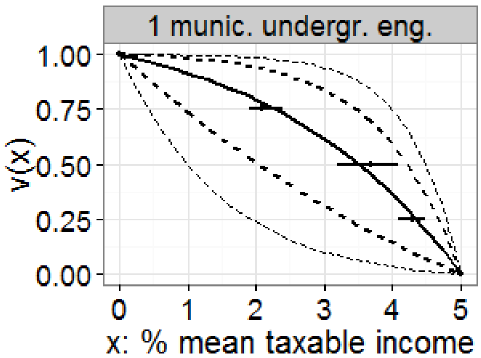


Figure SI4.12f: Elicited value function levels for 'Annual cost per inhabitant in% of the mean taxable income [%] (costcap)' and fitted distributions for one stakeholder. For detailed description see Fig. SI4.12a.

Marginal utility functions

Table SI4.5: Fitted marginal utility function curvature parameters (elicited certainty equivalents, and standard deviation of the fit not shown) for ten stakeholders. The bold numbers indicate parameter r_i, which determines the curvature of the function (see Material and methods, main text). These numbers were derived from fitting an exponential function to the elicited "best guess" and range of the certainty equivalent from the interview partner. *r*<0… risk seeking; *r*≈0… risk neutral; *r*>0… risk averse. "Overall" indicates the number of stakeholders assigned to one of three general forms of the utility function. "Summary" in the last row indicates for each stakeholder how many times one of the three respective forms was observed.

											Overall		
Attribute	**SH1**	**SH2**	**SH3**	**SH4**	**SH5**	**SH6**	**SH7**	**SH8**	**SH9**	**SH10**	r<0	r≈0	r>0
rehab	-	**2.95**	-	**0.97**	**2.78**	-	-	**2.18**	**-3.52**	-	4	0	1
adapt	-	-	-	-	-	-	-	-	-	-	0	0	0
gwhh	**-3.76**	**2.67**	**5.65**	**0.00**	-	-	**0.35**	**-0.06**	-	**r=0**	1	4	2
econs	-	-	-	-	-	-	-	-	-	-	0	0	0
No information about: vol_dw, reliab_dw													
vol_hw	-	-	-	-	-	-	-	-	-	-	0	0	0
reliab_hw	-	-	-	-	**r>0**	-	-	-	-	-	0	0	1
No information about: aes_dw, aes_hw													
faecal_dw	-	-	-	**-1.36**	**6.72**	**-52.79**	-	**-0.40**	-	-	1	1	2
No information about: cells_dw, faecal_hw, cells_hw, no3_dw, pest_dw, bta_dw, no3_hw, pest_hw, bta_hw, reliab_ffw, vol_ffw, efqm, voice, auton, time, area, collab													
costcap	**1.61**	-	-	-	-	-	-	-	-	-	0	0	1
costchange	-	-	**7.02**	**0.85**	-	-	-	-	-	**1.35**	0	0	3
SUM	r<0: 1 r≈0: 0 r>0: 1	r<0: 0 r≈0: 0 r>0: 2	r<0: 0 r≈0: 0 r>0: 2	r<0: 1 r≈0: 1 r>0: 2	r<0: 0 r≈0: 0 r>0: 3	r<0: 0 r≈0: 0 r>0: 1	r<0: 0 r≈0:0 r>0: 1	r<0 1 r≈0:2 r>0: 0	r<0: 1 r≈0: 0 r>0: 0	r<0: 0 r≈0: 1 r>0: 1	6	5	10

Acceptance thresholds

Table SI4.6: Expressed acceptance thresholds and potential preference interactions as stated by stakeholders. Comments for potential interactions were not considered in preference modeling as they were neither elicited in a structured manner, nor discussed with all stakeholders. However, they would affect the aggregation model as follows: if well performing values can compensate for badly performing values, an additive aggregation model is presumably appropriate, and a preference for balanced results is indicative for a non-additive aggregation model (e.g., the multiplicative, Cobb-Douglas, or mixed models).

	Acceptance thresholds(AT)	Potential interactions
SH1	None	Well performing values compensate badly performing values, but if all others perform… …badly: more risk prone. …well: more risk averse.
SH2	If the no. of days with hygienic concerns for drinking water >2 d/a, then the overall value is 0.	Balanced results are preferred to compensation of extremes. If all others perform badly: more risk averse.
SH3	none	-
SH4	If the no. of days with hygienic concerns for drinking water >2 d/a, then the overall value is 0. Additionally, if costs increase and drinking and/or household water quality do not meet the current regulation, then the overall value is 0.	If all others perform badly: more risk prone.
SH5	If the no. of days with hygienic concerns for drinking water >30 d/a, then the overall value is 0. Additionally, if the no. of days with water quantity restrictions is > 60 d/a, then overall value is 0.	If all others perform… …badly: risk prone. …well: risk averse.
SH6	If the no. of days with hygienic concerns of drinking water is higher than the current regulation, the overall value is 0.	
SH7	If either the no. of days with hygienic concerns of drinking water or the reliability of firefighting supply are worse than the current situation (status quo), then the overall value is 0.	Balanced results are preferred to compensation of extremes, but if all others perform… …badly: more risk prone
SH8	If the no. of days with hygienic concerns for drinking water >30 d/a, then the overall value is 0.	If all others perform badly: more risk averse.
SH9	If the no. of days with hygienic concerns for drinking water >0 d/a, then the overall value is 0.	Compensation between objectives is possible.
SH10	If more than 100 % of the natural groundwater recharge are utilized, then the overall value is 0. Additionally, if the no. of days with hygienic concerns or esthetic impairments for drinking water is >0 d/a, then the overall value is 0. A cost increase higher than 1% in 5 years is inacceptable; in that case, the overall value is 0.	High cost increases and high annual costs are not independent.

SI5 Uncertainty analysis

Table SI5.1: Mean μ and standard deviations σ of rank distributions of 11 alternatives (A1a–A9; see Tab. SI3.1) for four future scenarios and 10 stakeholders (SH; see Tab. SI1.1)

	SH1		SH2		SH3		SH4		SH5		SH6		SH7		SH8		SH9		SH10	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Boom																				
A1a	4.0	1.7	3.6	1.0	4.3	1.2	3.9	1.1	4.6	1.3	5.0	1.0	3.0	0.0	3.9	1.2	3.1	0.4	4.4	0.5
A1b	4.3	1.3	2.2	1.0	2.6	1.3	2.6	1.1	2.8	1.2	2.6	1.0	1.9	0.3	3.0	1.1	1.1	0.4	5.4	0.5
A2	4.4	1.8	1.7	0.9	2.8	1.3	2.3	1.2	1.6	0.9	1.4	0.7	1.1	0.3	2.8	1.3	2.0	0.4	3.1	0.6
A3	10.0	1.0	9.1	1.1	10.3	0.6	9.2	1.3	10.4	0.6	10.3	0.6	9.0	0.0	10.0	0.6	9.0	1.3	9.1	0.3
A4	9.1	0.9	8.1	1.2	9.2	0.7	7.8	1.4	10.1	0.8	10.0	0.7	10.0	0.0	9.2	0.8	8.8	1.3	8.1	0.2
A5	10.4	0.8	9.9	1.1	10.4	0.8	9.6	1.2	9.4	0.8	9.7	1.0	11.0	0.0	10.6	0.7	8.2	1.3	11.0	0.0
A6	8.3	0.8	5.5	1.5	7.8	1.0	6.5	1.3	5.7	2.5	4.8	1.9	5.3	1.1	7.7	0.8	4.5	1.2	9.9	0.5
A7	4.8	2.9	5.9	1.3	5.7	1.9	4.7	1.7	6.0	1.8	7.5	0.6	8.0	0.0	7.4	0.5	4.6	0.9	4.6	2.1
A8a	2.5	1.0	5.6	2.4	3.5	1.6	5.3	3.0	4.3	1.4	4.1	1.1	5.4	0.5	3.3	1.4	7.6	1.4	2.4	0.5
A8b	2.4	1.4	4.6	2.4	2.7	1.6	4.3	3.0	3.7	1.4	3.2	1.1	4.4	0.5	2.1	1.4	6.5	1.4	1.2	0.4
A9	5.6	1.1	9.8	1.5	6.7	0.7	9.8	1.6	7.5	0.7	7.4	0.7	6.9	0.3	6.0	0.2	10.5	1.0	6.9	0.4
Doom																				
A1a	4.8	2.2	4.9	1.8	3.8	1.5	4.7	2.3	4.8	1.5	4.7	1.0	3.5	0.5	3.9	0.3	4.6	2.0	7.4	0.7
A1b	3.7	2.3	3.1	2.0	1.6	1.0	3.6	2.3	2.4	1.3	2.2	0.7	1.8	0.4	2.1	0.4	2.9	2.3	6.4	0.7
A2	7.5	1.5	4.4	2.6	3.6	1.3	6.7	1.9	5.1	1.4	2.9	0.6	3.4	0.6	2.9	0.5	4.6	2.7	2.6	1.1
A3	10.4	1.9	8.6	2.3	10.8	0.6	8.6	2.5	10.9	0.3	10.8	0.4	9.6	0.5	10.5	0.7	9.0	2.2	10.1	0.8
A4	3.7	2.5	6.3	2.3	8.7	0.6	5.4	2.5	8.3	0.8	9.1	0.4	7.0	0.0	8.1	0.5	6.7	2.3	5.3	0.9
A5	9.9	1.1	8.6	2.2	10.1	0.6	8.1	2.4	9.9	0.8	10.0	0.8	11.0	0.0	10.4	0.5	7.4	2.2	10.2	0.9
A6	5.1	1.6	1.5	1.0	3.3	1.9	1.7	1.2	2.2	1.4	1.1	0.3	1.3	0.7	1.1	0.6	2.4	1.4	3.0	1.1
A7	5.5	3.7	5.6	2.4	6.3	2.2	4.6	2.7	5.9	2.3	7.3	0.5	9.4	0.5	8.9	0.6	3.5	1.7	9.7	0.6
A8a	4.0	1.1	5.4	0.9	5.1	1.4	4.9	0.9	4.2	1.5	5.0	0.9	5.0	0.0	5.0	0.1	6.2	0.7	1.7	0.9
A8b	3.1	1.2	7.1	1.7	4.7	1.5	7.2	1.9	3.7	1.7	5.1	0.8	6.0	0.0	6.0	0.1	7.8	1.3	2.8	0.9
A9	8.3	0.7	10.6	0.8	8.0	0.8	10.6	0.7	8.7	0.8	7.8	0.6	8.0	0.0	7.1	0.4	10.8	0.4	6.9	1.0
Quality of life																				
A1a	4.2	1.6	4.7	1.9	3.5	1.5	4.5	2.3	4.1	1.7	5.5	0.9	3.5	0.7	3.7	0.8	4.5	2.0	6.5	0.8
A1b	4.7	2.2	3.7	2.9	1.7	1.1	4.6	3.2	2.0	1.3	2.4	1.1	2.6	1.5	2.3	0.7	3.6	3.2	7.5	0.8
A2	7.5	1.7	4.0	1.6	4.1	1.3	6.7	1.0	5.3	1.3	3.3	1.0	2.9	0.7	2.8	0.8	3.7	1.7	3.3	0.6
A3	10.3	2.0	8.5	2.3	10.7	0.6	8.6	2.5	10.9	0.4	10.8	0.4	10.0	0.0	10.5	0.8	9.0	2.2	11.0	0.1
A4	4.1	2.4	7.0	2.3	9.0	0.5	5.8	2.5	8.5	0.8	9.2	0.4	8.0	0.1	8.7	0.6	6.8	2.3	5.4	0.9
A5	10.0	1.1	8.6	2.3	10.1	0.7	8.0	2.4	9.8	0.8	9.9	0.8	11.0	0.0	10.4	0.5	7.2	2.3	9.7	0.7
A6	6.0	1.5	1.6	1.2	3.8	1.9	2.4	1.9	3.3	1.7	1.2	0.5	1.5	1.0	1.3	1.0	2.4	1.5	3.6	0.6
A7	5.0	3.7	5.4	2.3	6.2	2.2	4.2	2.7	5.8	2.3	7.4	0.5	9.0	0.1	8.4	0.7	3.7	1.7	9.1	0.4
A8a	3.6	1.2	5.2	1.2	5.0	1.6	4.3	1.3	4.4	1.6	4.8	1.0	4.5	1.1	5.9	0.4	6.4	1.0	1.1	0.3
A8b	2.4	1.2	6.9	1.8	4.0	1.6	6.3	2.6	3.4	1.6	3.8	1.0	6.0	0.1	4.9	0.4	7.9	1.3	2.0	0.4
A9	8.2	0.9	10.5	1.0	7.9	0.7	10.6	0.7	8.6	0.8	7.7	0.6	7.0	0.0	7.0	0.2	10.8	0.4	6.8	1.1
Status quo																				
A1a	3.6	2.3	4.4	1.5	3.5	1.5	4.2	2.0	4.2	1.6	4.5	0.9	3.3	0.6	3.9	0.3	4.5	1.6	1.9	0.8
A1b	4.7	2.1	2.7	1.8	1.6	1.0	3.2	2.0	1.9	1.2	2.1	0.6	1.7	0.4	2.1	0.5	2.7	2.0	1.4	0.7
A2	7.5	1.5	4.3	2.1	3.7	1.3	6.9	1.7	5.3	1.3	2.9	0.5	3.5	0.6	2.8	0.5	4.2	2.4	4.8	1.1
A3	10.3	2.0	8.6	2.3	10.7	0.7	8.6	2.5	10.8	0.4	10.8	0.4	10.0	0.0	10.5	0.7	9.0	2.2	9.1	0.3
A4	3.5	2.2	6.3	2.3	8.7	0.7	5.3	2.5	8.3	0.9	9.1	0.5	7.0	0.0	8.1	0.5	6.7	2.3	7.0	0.6
A5	9.9	1.2	8.6	2.2	10.1	0.7	8.1	2.4	9.9	0.8	9.9	0.8	11.0	0.0	10.4	0.5	7.4	2.3	10.8	0.4
A6	5.8	1.5	1.5	1.1	3.6	1.9	2.2	1.6	2.8	1.6	1.1	0.3	1.4	0.8	1.2	0.7	2.3	1.4	5.0	1.2
A7	5.2	3.7	5.6	2.3	6.2	2.3	4.4	2.7	5.8	2.3	7.4	0.6	9.0	0.0	8.9	0.6	3.6	1.7	10.0	0.5
A8a	2.9	1.0	6.1	1.6	4.8	1.4	5.6	1.8	3.6	1.5	4.8	0.6	5.0	0.0	5.3	0.5	6.8	1.3	3.4	1.1
A8b	4.1	1.0	7.1	1.7	4.9	1.6	7.0	2.1	4.6	1.5	5.6	0.7	6.0	0.0	5.7	0.5	7.9	1.3	4.4	1.1
A9	8.5	0.7	10.7	0.7	8.1	0.8	10.6	0.7	8.8	0.7	7.8	0.6	8.0	0.0	7.1	0.4	10.8	0.4	7.9	0.3

Table SI5.2: Median rank (MR) and interquartile ranges (IQR) of rank distributions of 11 alternatives (A1a–A9; see Tab. SI3.1) for four future scenarios and 10 stakeholders (SH; see Tab. SI1.1)

	SH1		SH2		SH3		SH4		SH5		SH6		SH7		SH8		SH9		SH10	
	MR	IQR	MR	IQR	MR	IQR	MR	IQR	MR	IQR	MR	IQR	MR	IQR	MR	IQR	MR	IQR	MR	IQR
Boom																				
A1a	4	2	3	2	5	2	3	2	5	3	5	1	3	0	4	2	3	0	4	1
A1b	5	2	2	0	3	3	2	2	2	2	2	1	2	0	3	2	1	0	5	1
A2	5	3	1	1	3	2	2	2	1	1	1	1	1	0	3	1	2	0	3	0
A3	10	1	9	1	10	1	9	1	11	1	10	1	9	0	10	0	9	2	9	0
A4	9	2	8	1	9	1	8	2	10	2	10	2	10	0	9	1	9	2	8	0
A5	11	1	10	1	11	1	10	2	9	1	9	2	11	0	11	0	8	1	11	0
A6	8	1	6	3	8	0	7	2	7	4	6	3	6	2	8	1	4	1	10	0
A7	7	6	6	2	6	1	5	2	6	1	8	1	8	0	7	1	5	1	6	4
A8a	3	1	5	3	3	3	5	6	4	2	4	1	5	1	2	3	7	1	2	1
A8b	2	2	4	3	2	3	4	6	4	2	3	1	4	1	1	3	6	1	1	0
A9	6	1	11	3	7	1	11	3	8	1	7	1	7	0	6	0	11	0	7	0
Doom																				
A1a	5	5	4	2	3	2	4	4	5	2	4	2	3	1	4	0	4	3	8	1
A1b	4	5	2	2	1	1	3	4	2	2	2	1	2	0	2	0	2	4	7	1
A2	8	3	3	3	3	1	7	4	5	2	3	0	3	1	3	0	4	5	3	1
A3	11	0	9	1	11	0	10	5	11	0	11	0	10	1	11	1	10	1	10	2
A4	2	3	7	5	9	1	7	5	8	1	9	0	7	0	8	0	8	4	5	0
A5	10	0	9	2	10	0	9	5	10	0	10	0	11	0	10	1	9	4	11	2
A6	6	3	1	0	3	4	1	1	2	2	1	0	1	0	1	0	2	3	3	2
A7	7	8	7	4	7	3	3	5	7	2	7	1	9	1	9	0	3	3	10	1
A8a	4	1	5	1	5	2	5	2	4	2	5	1	5	0	5	0	6	0	1	2
A8b	3	2	6	2	5	2	7	3	4	3	5	1	6	0	6	0	7	2	2	2
A9	8	1	11	0	8	2	11	1	9	1	8	1	8	0	7	0	11	0	6	2
Quality of life																				
A1a	4	2	4	3	3	2	4	5	4	3	6	2	4	1	4	1	4	3	6	1
A1b	5	3	2	4	1	1	4	7	1	2	2	1	2	3	2	1	2	6	8	1
A2	8	3	3	2	4	2	7	1	6	2	3	1	3	1	3	1	3	3	3	1
A3	11	0	9	2	11	0	10	5	11	0	11	0	10	0	11	1	10	1	11	0
A4	4	4	8	2	9	0	7	5	8	1	9	0	8	0	9	1	8	5	5	0
A5	10	0	9	2	10	0	9	5	10	0	10	0	11	0	10	1	8	5	10	0
A6	6	2	1	0	4	4	1	4	3	3	1	0	1	1	1	0	1.5	3	4	1
A7	6	8	7	5	7	3	3	5	7	2	7	1	9	0	8	1	5	3	9	0
A8a	4	2	5	1	5	2	4	3	4	3	5	2	5	0	6	0	6	0	1	0
A8b	3	2	6	2	4	2	5	5	3	3	4	2	6	0	5	0	7	2	2	0
A9	8	1	11	1	8	1	11	1	9	1	8	1	7	0	7	0	11	0	6	2
Status quo																				
A1a	4	5	4	2	3	3	4	4	4	3	4	2	3	1	4	0	4	3	2	1
A1b	5	4	2	2	1	1	3	4	1	1	2	0	2	1	2	0	2	4	1	1
A2	8	2	4	3	4	2	7	3	6	2	3	0	4	1	3	0	3	5	5	2
A3	11	0	9	1	11	0	10	5	11	0	11	0	10	0	11	1	10	1	9	0
A4	2	3	7	5	9	1	7	5	8	1	9	0	7	0	8	0	8	4	7	0
A5	10	0	9	2	10	0	9	5	10	0	10	0	11	0	10	1	8	5	11	0
A6	6	2	1	0	4	4	1	2	2	3	1	0	1	1	1	0	2	3	5	2
A7	6	8	7	4	7	3	3	5	7	2	7	1	9	0	9	0	4	3	10	0
A8a	3	1	5	2	5	2	5	3	4	3	5	0	5	0	5	1	6	2	3	2
A8b	4	1	6	2	5	2	6	4	5	3	6	1	6	0	6	1	7	2	4	1
A9	9	1	11	0	8	1	11	1	9	1	8	1	8	0	7	0	11	0	8	0

Table SI5.3: Difference between rankings based on usual simplified assumptions and median ranking based on uncertain preferences. μ_{SH1-10} = mean, ∑|x| = sum of the absolute rank differences. "0": equal rank, negative (positive) values indicate a ranking which is worse (better) under simplifying assumptions.

	Difference between rankings using simplified assumptions or uncertain preferences										
	SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	μ_{SH1-10}
Boom											
A1a	-1	0	0	1	2	2	0	1	0	-1	0.8
A1b	-1	0	1	1	0	0	0	1	0	-1	0.5
A2	1	0	2	-1	0	0	0	2	0	0	0.6
A3	0	0	0	0	1	0	0	0	0	0	0.1
A4	0	0	1	1	1	1	0	1	1	0	0.6
A5	0	0	0	0	-2	-2	0	0	-2	0	0.6
A6	0	-1	-1	-1	0	-2	0	-1	-3	0	0.9
A7	0	0	-1	-1	-2	1	0	0	-1	2	0.8
A8a	1	0	-1	0	-1	-1	0	-3	2	0	0.9
A8b	1	0	-1	0	0	-1	0	-3	2	0	0.8
A9	3	0	1	0	2	1	0	0	0	0	0.7
∑|x|	**8**	**1**	**9**	**6**	**11**	**11**	**0**	**12**	**11**	**4**	**7.3**
Doom											
A1a	2	0	0	1	2	-2	0	0	2	0	0.9
A1b	2	0	0	1	0	0	0	-1	1	0	0.5
A2	4	0	1	2	1	-2	-1	1	1	2	1.5
A3	1	0	1	1	1	1	1	1	1	-1	0.9
A4	-6	-1	0	1	0	0	0	1	0	-1	1
A5	-1	-1	-1	-1	-1	-1	0	-1	-1	1	0.9
A6	5	0	-1	0	1	0	0	0	-2	1	1
A7	-2	2	-1	-5	-2	0	-1	0	-4	1	1.8
A8a	-2	-1	-1	1	-2	2	0	0	1	-3	1.3
A8b	-2	-1	0	0	-1	1	0	0	1	-1	0.7
A9	1	0	1	0	2	0	0	-1	0	1	0.6
∑|x|	**28**	**6**	**7**	**13**	**13**	**9**	**3**	**6**	**14**	**12**	**11.1**
Quality of life											
A1a	3	0	1	2	1	0	1	0	2	-1	1.1
A1b	3	0	0	1	0	0	0	-1	1	0	0.6
A2	2	0	1	1	2	0	-1	1	0	0	0.8
A3	1	0	1	1	1	1	0	1	1	0	0.7
A4	-4	0	0	-1	0	0	0	1	0	-1	0.7
A5	-1	-1	-1	-1	-1	-1	0	-1	-2	0	0.9
A6	1	0	0	0	1	0	0	0	-2.5	0	0.45
A7	-3	2	-1	-4	-2	0	0	-1	-2	0	1.5
A8a	0	-1	-1	0	-2	0	0	0	1	0	0.5
A8b	0	-1	-1	0	-2	0	0	0	1	0	0.5
A9	1	0	1	0	2	0	0	0	0	1	0.5
∑|x|	**19**	**5**	**8**	**11**	**14**	**2**	**2**	**6**	**12.5**	**3**	**8.25**
Status quo											
A1a	3	0	1	1	1	-2	0	0	2	1	1.1
A1b	3	0	0	1	-1	0	0	0	1	-1	0.7
A2	2	1	1	3	2	0	0	0	0	2	1.1
A3	1	0	1	1	1	1	0	1	1	-1	0.8
A4	-5	-1	0	1	0	0	0	1	0	-1	0.9
A5	-1	-1	-1	-1	-1	-1	0	-1	-2	0	0.9
A6	3	0	0	0	1	0	0	0	-2	1	0.7
A7	-3	2	-1	-5	-2	0	0	0	-3	1	1.7
A8a	-1	-1	-1	0	-1	0	0	0	1	-2	0.7
A8b	-1	-1	0	-1	-1	2	0	0	1	-2	0.9
A9	1	0	1	0	2	0	0	-1	0	1	0.6
∑|x|	**24**	**7**	**7**	**14**	**13**	**6**	**0**	**4**	**13**	**13**	**10.1**

SI6 Global sensitivity analysis (GSA)

Stability of the ranking of alternatives to attribute sample size n

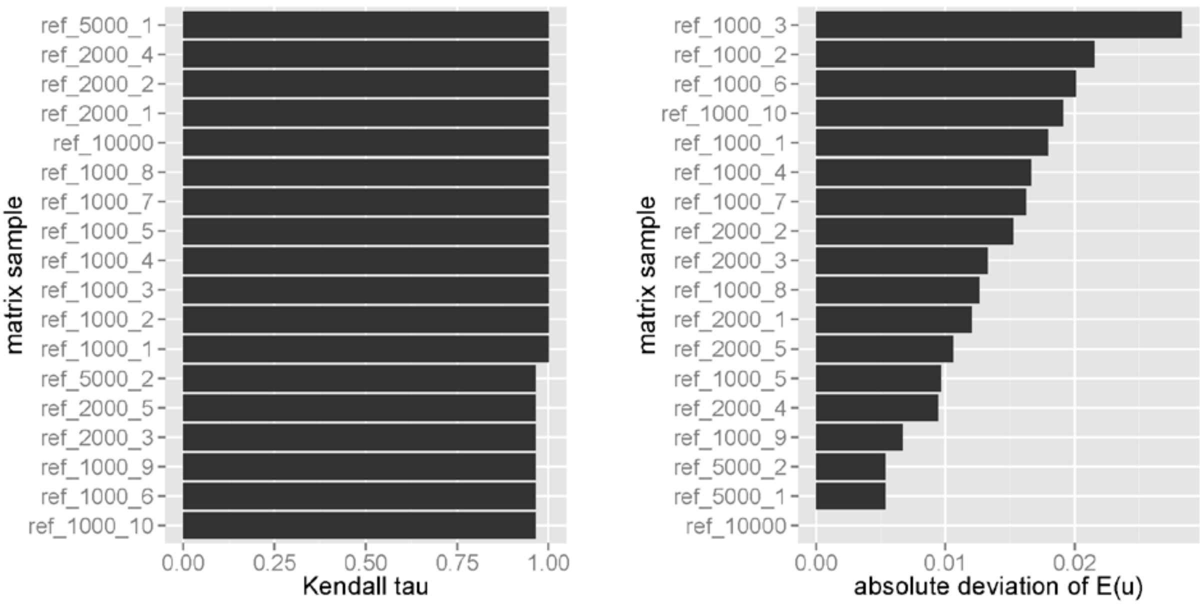


Figure SI6.1: Change of Kendall rank correlation coefficient τ (left) and expected utility E(u) (right) depending on the size and portion of the underlying attribute sample. ref_10000 represents the reference using the overall n=10'000 attribute sample. The ranking using this sample was compared to sub-samples of different sizes, n=5'000, n=2'000, and n=1'000. While the deviation of the expected utility E(u) is strongly dependent on the sample size n (right), the ranking is not (left).

Stability of sensitivity coefficients to preference parameter sample size s

Table SI6.1: Top 15 parameters ranked by first order index (main effect) for different preference parameter samples s. Preferences of SH2, the reference ranking is based on the mean parameters E(**θ**) and attribute predictions for the Status quo scenario.

	s = 8000				s = 4000				n = 2000			
par	**rank first**	**first order**	**total order**	**rank total**	**rank**	**first order**	**total order**	**rank total**	**rank**	**first order**	**total order**	**rank total**
r	1	0.7226	0.9145	1	**1**	0.7169	0.9099	1	1	0.6951	0.9080	1
a.IE	2	0.0206	0.0992	3	**2**	0.0190	0.0780	5	2	0.0197	0.0912	3
a.overall	3	0.0105	0.1393	2	**3**	0.0097	0.1242	3	3	0.0098	0.1314	2
c.IE_rehab	4	0.0098	0.0650	5	**4**	0.0093	0.0436	7	4	0.0088	0.0542	5
a.SA	5	0.0040	0.0590	6	**5**	0.0034	0.0369	2	5	0.0042	0.0478	7
a.WS_dw	6	0.0027	0.0532	8	**6**	0.0021	0.0267	11	6	0.0035	0.0482	6
c.IE_flex	7	0.0009	0.0929	4	**7**	0.0016	0.0722	12	7	0.0017	0.0865	4
c.WS_dw.reliab	8	0.0008	0.0464	13	**8**	0.0011	0.0240	6	9	0.0010	0.0409	9
c.WS_ffw.quant	9	0.0008	0.0471	10	**9**	0.0008	0.0212	8	8	0.0011	0.0377	14
c.RG_energ	13	0.0005	0.0553	7	**10**	0.0008	0.0306	4	10	0.0010	0.0466	8
c.SA_time	10	0.0006	0.0452	18	**11**	0.0008	0.0192	78	13	0.0006	0.0358	26
c.SA_auton	14	0.0004	0.0463	14	**12**	0.0006	0.0214	10	11	0.0009	0.0382	12
c.SA_efqm	15	0.0004	0.0477	9	**13**	0.0005	0.0228	30	12	0.0008	0.0387	11
c.WS_ffw.reliab	11	0.0005	0.0454	17	**14**	0.0004	0.0220	16	14	0.0006	0.0373	16
w2	18	0.0002	0.0397	78	**15**	0.0003	0.0178	47	15	0.0005	0.0255	79
$\sum\theta_z$		0.7789	4.8907			0.7709	2.7483			0.7572	3.9912	
$\sum w_i$		0.0021	1.7713			0.0021	0.7598			0.0043	1.2466	
$\sum c_j$		0.0157	1.4024			0.0172	0.6669			0.0192	1.1753	
$\sum\alpha_k$		0.0385	0.8025			0.0346	0.4117			0.0386	0.6613	

Figure SI6.2: First and total order sensitivity coefficients for three different sample sizes s. Stakeholder SH2, Status quo scenario. *r* is the overall risk attitude, parameters starting with "*a.*" are the aggregation mixture parameters, "*c.*" value function curvature parameters, and "*w.*" the weighting parameters. Parameter names begin with the parameter group ("*a.*" or "*c.*"), followed by the respective main objective of the branches going down the hierarchy up to the indicated end point (see Fig. 1) in main text), i.e. aggregation node or attribute. Acronyms for the top-level main objectives (and weight numbers, more details see Tab. SI4.1) are: "IE" – "*high intergenerational equity* (*w.1*)", "RG" – "*high resources and groundwater protection (w.2)*", "WS" – "*good water supply (w.3)*", "SA" – "*high social acceptance (w.4)*", and "KO" – "*low costs (w.5)*". E.g. "*c.WS_dw.reliab*" stands for the value function curvature of the objective "*high reliability (reliab)*" of the drinking water supply (WS_dw). "*a.overall*" – mixture parameter at the highest hierarchical level.