DISS. ETH NO. 21467

Multi-criteria decision analysis for water supply infrastructure planning under uncertainty

A dissertation submitted to

ETH ZURICH

for the degree of

Doctor of Sciences

presented by

Lisa Scholten

Dipl.-Ing. Wasserwirtschaft, Technische Universität Dresden
born September 28, 1984
citizen of Germany

accepted on the recommendation of

Prof. Dr. Peter Reichert, examiner
Dr. Judit Lienert, co-examiner
Prof. Dr. Max Maurer, co-examiner
Prof. Dr. Alec Morton, co-examiner

2013

Summary

Our centralized water supply systems are aging. Despite their success in reliably providing high quality drinking water, nowadays especially small utilities (e.g. less than 10'000 inhabitants supplied) are ill-prepared to face possible future challenges. The fragmented structure of the water supply sector leads to a lack of institutional, financial and personnel resources for professional management and planning of water supply systems. Current planning is furthermore challenged by insufficient knowledge and data about the prevailing water infrastructure condition and future rehabilitation demand. It usually ignores future dynamics and planning uncertainties, as well as alternatives to the perpetuation of the status quo. Infrastructure decision making is usually not transparent and only few stakeholders are included into the decision process.

This thesis presents an approach to overcome these shortcomings and support long-term water supply infrastructure planning under uncertainty in a multi-stakeholder context. Thereto, methods from multi-criteria decision analysis (MCDA), strategic asset management (SAM), pipe failure modeling, and scenario planning were combined, adapted, and further developed. The suitability of the approach was validated in a case study in Switzerland.

To improve the prediction of pipe service life in view of data scarcity, it was shown how the knowledge of experts can be quantitatively assessed and integrated into the calibration of pipe survival models by means of Bayesian inference. Similarly, knowledge gained from three mid-size to large Swiss water networks was used to improve calibration of a novel pipe failure model. It is demonstrated that this failure model is able to deal with the common data situation, and mitigate overestimation of the time to failure caused by the absence of data from already replaced pipes. The failure model was combined with a rehabilitation model to assess the performance of 18 rehabilitation strategies under four future scenarios for a small water utility. MCDA was used to compare these alternatives under different preferences concerning three objectives. The analysis revealed that the common strategy, purely reactive rehabilitation, is not recommendable in most cases and that annual replacement of 1–2 % of the network by condition might be a good strategy for the utility in question.

These findings were considered during the definition of alternatives for a second MCDA study aiming at identifying 'good water supply infrastructure' alternatives. Eleven alternatives and an objectives hierarchy consisting of 44 fundamental objectives and 30 attributes were developed together with stakeholders. The alternatives differ not only with regard to rehabilitation management, but also technical, managerial, and organizational aspects. The outcomes of all alternatives regarding these attributes were predicted under four future scenarios to account for uncertainties about the future development. The approach for the elicitation and modeling of preferences includes the imprecision of the stated preferences as well as uncertainties of preference parameters which were not elicited (the aggregation model, marginal value functions, risk attitude, and scaling factors). Preferences of ten selected stakeholders were then elicited and probability distributions of the ranking of alternatives based on these preferences were obtained. Despite differences in the individual rankings, a potential compromise solution could be proposed and ways for potential adaptation and improvement of other alternatives be indicated. In general, alternatives with good outcomes regarding groundwater protection, water quality, supply reliability, and realization of the rehabilitation demand received the highest ranks, as these were also among the most important objectives for the majority of the stakeholders. As operation and management do considerably contribute to the performance regarding the latter three objectives, the importance of a thorough infrastructure and rehabilitation management cannot be neglected. In view of the possible ranges of the outcomes, the objectives of 'high social acceptance' (e.g. disturbance by unnecessary road works, resources autonomy), and to some extent also 'low costs' were judged less important.

Multi-criteria decision analysis proved useful to support the long-term planning of water infrastructures under uncertainty and in a multi-stakeholder framework. The usual extrapolation of the status quo was overcome. Future dynamics and uncertainties could be incorporated by combining decision making and modeling with scenario planning, besides the quantitative consideration of uncertainties in making predictions, and evaluating the results. With the presented approaches for the modeling of pipe failures and rehabilitation, the methods and tools for the assessment of the current condition and future rehabilitation demand of small water networks despite a difficult data situation are now available.

Zusammenfassung

Unsere zentralen Wasserversorgungssysteme altern. Trotz ihres Erfolgs im Hinblick auf eine zuverlässige Versorgung mit Trinkwasser hoher Qualität, sind heute vor allem kleine Wasserversorgungen (weniger als 10'000 versorgte Einwohner) schlecht auf mögliche zukünftige Herausforderungen vorbereitet. Die fragmentierte Struktur des Wasserversorgungssektors führt zu einem Mangel institutioneller, finanzieller und personeller Ressourcen für ein professionelles Management und die langfristige Planung der Wasserversorgung. Die derzeitige Planung steht weiterhin vor der Herausforderung mangelnden Wissens und Daten sowohl über den Zustand, als auch den zukünftigen Erneuerungsbedarf der Infrastrukturen. Üblicherweise werden weder Planungsunsicherheiten oder die zukünftige Dynamik, noch Alternativen zur Fortführung des Status quo berücksichtigt. Die Entscheidungsfindung in Infrastrukturprojekten ist häufig intransparent und nur wenige Interessensgruppen sind an dem Entscheidungsprozess beteiligt.

Diese Dissertation präsentiert eine Herangehensweise zur Überwindung dieser Defizite und zur Unterstützung von Entscheidungsprozessen in der langfristigen Wasserinfrastrukturplanung unter Miteinbeziehung verschiedener Akteure und unter Berücksichtigung von Unsicherheiten. Dazu werden Methoden der multi-kriteriellen Entscheidungsanalyse (MCDA), des strategischen Asset-Managements (SAM) und der Schadensmodellierung von Wasserleitungen kombiniert, angepasst und weiterentwickelt. Die Validierung erfolgte in einer Fallstudie in der Schweiz.

Zur Verbesserung der Nutzungsdauerprognose von Wasserleitungen vor dem Hintergrund mangelnder Daten wird veranschaulicht, wie Expertenwissen quantitativ erhoben und in die Kalibrierung von Überlebensmodellen mittels Bayesscher Inferenz einbezogen werden kann. Auf ähnliche Weise werden Erkenntnisse, die basierend auf drei mittelgrossen bis grossen Wassernetzen gewonnen wurden, für die Verbesserung der Kalibrierung eines neuartigen Leitungsausfallmodells verwendet. Es wird gezeigt, dass dieses Modell in der Lage ist, mit der allgemeinen Datensituation umzugehen und die durch fehlende Daten bereits ersetzter Leitungen hervorgerufene Überschätzung der Zeit zu einem Schaden zu verringern. Durch Kombination dieses Modells mit einem Rehabilitationsmodell wurde die Leistung von 18 Rehabilitationsstrategien unter vier Zukunftsszenarien für eine kleine Wasserversorgung ermittelt. Für den Vergleich dieser Alternativen anhand der Präferenzen für drei Ziele wird die MCDA vorgestellt und angewandt. Die Analyse zeigte, dass der verbreitete, rein reaktive Ersatz von Wasserleitungen in den meisten Fällen nicht zu empfehlen ist und dass ein zustandsbedingter jährlicher Ersatz von 1–2 % des Netzes eine gute Strategie für das betroffene Wasserversorgungsunternehmen darstellt.

Diese Erkenntnisse flossen in die Definition von Alternativen in einer zweiten MCDA-Studie ein, die zum Ziel hatte, Optionen für eine ‚gute Wasserversorgungsinfrastruktur' zu identifizieren und einen robusten Umgang mit Unsicherheiten aufzuzeigen. Elf Alternativen und eine aus 44 Zielen und 30 Attributen bestehende Zielhierarchie wurden hierfür gemeinsam mit Interessensvertretern entwickelt. Die Alternativen unterscheiden sich nicht nur im Hinblick auf Eigenschaften wie das Rehabilitationsmanagement, sondern auch technische, betriebliche und organisatorische Aspekte, die mitunter sehr verschieden zum heutigen System sind. Um Unsicherheiten über die zukünftige Entwicklung zu berücksichtigen, wurde die Leistung der Alternativen im Sinne der Szenarienplanung unter vier (ebenfalls mit Akteuren entwickelten) Szenarien vorhergesagt. Der Ansatz zur Erhebung und Modellierung der Präferenzen berücksichtigt sowohl die Ungenauigkeit der Präferenzangaben, als auch Unsicherheiten von Präferenzparametern, die nicht erhoben wurden (das Aggregationsmodel, die marginalen Wertefunktionen, die Risikoeinstellung und die Skalierungsfaktoren). Von zehn ausgewählten

Interessensvertretern wurden ungenaue Präferenzen erhoben und anhand dieser Präferenzen Wahrscheinlichkeitsverteilungen der Rangierung der Alternativen generiert. Trotz individueller Unterschiede konnten eine potenzielle Kompromisslösung vorgeschlagen und Wege für die Anpassung und Verbesserung von Alternativen aufgezeigt werden. Insgesamt schnitten Alternativen mit guten Ergebnissen bezüglich Grundwasserschutz, Wasserqualität, Zuverlässigkeit der Versorgung und Umsetzung des Rehabilitationsbedarfs am besten ab, da diese Ziele von der Mehrheit der Akteure als sehr wichtig eingeschätzt wurden. Da der Betrieb und Unterhalt die Leistung bezüglich der letzten drei Ziele massgeblich beeinflussen, ist die Bedeutung eines sorgfältigen Infrastruktur- und Rehabilitationsmanagements nicht zu vernachlässigen. Im Hinblick auf die möglichen Ausprägungen der Ergebnisse waren die Ziele 'hohe soziale Akzeptanz (z.B. Störung durch unnötige Strassenarbeiten, Ressourcenautonomie) und teilweise auch ‚niedrige Kosten' weniger relevant.

Die MCDA stellte sich als nützlich für die langfristige Wasserinfrastrukturplanung unter Unsicherheit und Miteinbeziehung verschiedener Akteure und Präferenzen heraus. Die übliche Extrapolation des Status quo wurde überwunden. Mittels der Kombination von MCDA und Szenarienplanung, sowie der quantitativen Berücksichtigung von Unsicherheiten beim Aufstellen von Prognosen und Evaluieren der Resultate konnten Unsicherheiten unterschiedlicher Quellen berücksichtigt werden. Mittels der vorgestellten Herangehensweisen für die Schadens- und Rehabilitationsmodellierung sind nun Methoden und Werkzeuge vorhanden, mittels derer der aktuelle Zustand und der zukünftige Rehabilitationsbedarf kleiner Wasserversorgungen trotz schwieriger Datenlage bestimmt werden können.

Contents

1 Introduction

1.1 Background

A continuous supply with drinking water is a commodity that is taken for granted in highly developed countries. Past generations built extensive networks of water supply facilities permitting outstanding levels of service. Apart from its paramount importance for human wellbeing and economic productivity, it also constitutes one of the largest public capital investments in history. In Switzerland, the replacement value is estimated at 50 billion Swiss francs (Martin, 2009), i.e. roughly 6'300 CHF per inhabitant (ca. €5'100 / US$6'800 at the time of writing). Although approximately 60–80 % thereof falls upon the underground pipe network alone (which also consumes about the same portion in water supply-related annual operation and maintenance expenditures; SVGW, 2009), little is known about the actual condition and the rehabilitation needs of this underground pipe infrastructure. Recent reports assume that about half of the Swiss public water supply infrastructure has reached its useful life (Martin, 2009; Swissplan, 2012) and that more reinvestments are necessary to keep the current levels of service, a problematic also recognized e.g. in Australia (Burns et al., 1999), and the USA (Selvakumar and Tafuri, 2012). This leads to questions such as when, where, and how to rehabilitate the current system, but also opens a window of opportunity for evaluating the well-served system against other (potentially more sustainable) options available today.

The main planning instrument in Switzerland is the "General Water Supply Project" (in German: "Generelles Wasserversorgungsprojekt", GWP, e.g. AfU, 2006; AUE, 2012; AWA, 2011; AWEL, 2010), mandatory by law in some of the cantons (Eawag, 2009). Its central task is to regularly assess the system's condition and rehabilitation demand in about 10 to 15-yearly intervals and to plan system adaptations on the long-term. As most of the Swiss water utilities are small in size, and legislation delegates water supply responsibilities to the municipalities, water infrastructure planning is usually performed on a very local scale.

1.2 Motivation

Long-term water infrastructure planning in Switzerland is facing a number of challenges: 1) limited institutional, financial, and professional capacities of the utilities, 2) little knowledge about water infrastructure condition and rehabilitation demand today and in the future, 3) narrow-minded extrapolation of status quo and negligence of future dynamics and uncertainty in planning, 4) lack of multi-stakeholder and transparent decision support.

Some of these challenges can be attributed to the small utility size and high fragmentation of the Swiss water sector (Dominguez et al., 2009; Lienert et al., 2013). More than half of the population is supplied by utilities of less than 10'000 serviced inhabitants (SVGW, 2009). In many places there are several water utilities providing water services for one municipality. In 2010, approximately 3'000 water utilities (Eawag, 2009) provided water supply for about 2'600 municipalities (BFS, 2010). Understaffing and restricted budgets are common in small utilities (Eawag, 2009), resulting in a lack of personnel and other resources for professional management – let alone thorough planning – of the water supply infrastructure. As a consequence, management and rehabilitation follow short-sighted, predominantly reactive strategies, which are rarely based on an informed analysis of the specific system conditions.

Besides this, the assessment and reliable prediction of rehabilitation demand is often impeded by a lack or an inadequate quality of pipe network data, not only in Switzerland (Alegre, 2010; U.S. Government Accountability Office, 2004). This lack concerns information about pipe characteristics such as age, material, or bedding on the one hand, and records about current pipe condition or past failures and rehabilitation measures on the other (Kleiner and Rajani, 1999; Le Gat, 2009; Renaud et al., 2009; Renaud et al., 2012). Pipe characteristics are important to prioritize pipes and ensure the efficient management of resources. At the same time, pipe condition data are needed to estimate the future condition, related performance (e.g. regarding water quality, service pressure, reliability, water losses, costs), and rehabilitation demand. Without these data, chances for a more proactive management of the existing assets are limited. Another difficulty is that existing failure and rehabilitation models are data intensive, such that their meaningful application is limited even with perfect documentation, due to few data given small network sizes.

In addition to this, and despite the long-term intentions of the GWP, current strategic planning in the Swiss water sector tends to ignore broader goals, large context uncertainties, and alternative solutions (Störmer et al., 2009). This issue is not particular to Switzerland, but can be considered a general challenge in current water infrastructure planning (e.g. see also Ferguson et al., 2013). On the same note, Ashley et al. (2008) report that *"[…] current governance, institutional, legislative, regulatory, risk, technological, and economic paradigms [in the UK] tend to constrain most water service providers into adopting well-tried and tested technologies"*. Possible drivers of change that encourage not only the consideration of alternatives to current technology and management, but also their evaluation under different future scenarios are: e.g. population dynamics, socio-economic development, water availability, water demand, and changes in regulation. This reliance on "business as usual" is not sustainable and, as Gander (2009) states with regard to climate change effects, the longevity of water supply facilities makes it necessary to adapt early to the emerging changes.

The strong link between technologies and institutions in Switzerland (Dominguez et al., 2009; Störmer et al., 2009), might further consolidate this situation and results in narrow-minded and exclusive, if not intransparent decision making. In a highly fragmented system where many different stakeholders interact (Lienert et al., 2013), it appears logical that these stakes and perspectives should be involved into decision making. Actual planning, however, has a long tradition of exclusive collaboration between the municipality and / or utility with local engineering companies. Besides the many positive aspects of this collaboration, innovation in infrastructure planning and the inclusion of further important stakeholders is unlikely to happen unless externally triggered. A precondition for this innovation are tools and methodologies to better support the strategic infrastructure planning process, as well as to assess the performance of the current system and alternatives based on the available information.

1.3 State of research and proposed approach

The main objective of the research presented in this PhD thesis is to develop a methodology for more comprehensive and integrative long-term water supply infrastructure planning. This methodology shall improve planning by making water supply alternatives more comparable, be inclusive of different stakes, and deal with lacking data and uncertain future developments. In particular, three research questions will be addressed:

(1) How can the decision makers' preferences be included into water infrastructure decision support?

(2) How can consequences of different water infrastructure alternatives be evaluated in face of an uncertain future?

(3) Which are optimal management strategies of water supply infrastructure under different change scenarios for a specific case study in Switzerland?

Two decision problems are addressed herein. The first concerns water supply planning and the fundamental question of how a 'good water supply infrastructure' can be achieved in a sustainable manner by different technical, organizational, and managerial options (Scholten et al., submitted; chapter 6). Since the long-term maintenance and rehabilitation management of the aging water networks is a major issue for utility managers in today's water infrastructure management, it will constitute the second, more specific decision problem. Here, an MCDA in search of 'good long-term rehabilitation strategies' will be carried out (see Scholten et al., 2014; chapter 5).

Water supply infrastructure planning

The existing water supply infrastructures were built to ensure a continuous supply of drinking water to increase the health and wellbeing of citizens, to provide fire security, and to provide water as a resource for economic activity. More recently, these water supply infrastructures are increasingly criticized for their limited ability to cope with a range of challenges, namely the dynamically changing socio-economic and socio-political environments, urbanization, and of course climate and environmental change (Ferguson et al., 2013; Ruth et al., 2007; Sharma et al., 2010). As a result, the consideration of sustainability as long-term objective in water supply planning, e.g. when transitioning to *water sensitive cities* (Brown et al., 2009; Ferguson et al., 2013), has gained increased attention. The high inflexibility of centralized pipe systems, increasing water scarcity and water losses, and the need for rehabilitation of deteriorating infrastructures, have led to the suggestion of more fit-for-purpose supply approaches or even decentralized water supply systems to increase sustainability (e.g. Sharma et al., 2010; Wong and Brown, 2009). A collection of alternative water supply technologies, covering point-of-use treatment, water recycling in households, and rainwater harvesting, is given in e.g. Makropoulos and Butler (2010). Other technical alternatives aim at reducing the (over-) dimensioning of the current pipe network to avoid hygienic and esthetic impairment of water quality, besides optimizing costs (Vreeburg et al., 2009). Additionally, alternative forms of utility governance, e.g. the regionalization and / or partial to full privatization of water utilities are sometimes considered to increase professionalism and efficiency (Dominguez et al., 2009; Lieberherr et al., 2012).

In contrast to these mostly technological considerations, Brown and Farrelly (2009) found that the most important barriers to sustainable urban water management are of socio-institutional nature. According to their study, more than 40 % out of an overall of 53 studies from Australia, New Zealand and Canada, identified a lack of coordination in the institutional framework. Other important barriers were: limited community engagement, empowerment, and participation, retards due to the regulatory framework, insufficiency of human and capital resources, unclear and fragmented responsibilities, poor organizational commitment, and a lack of information and understanding in applying integrated, adaptive forms of management, the absence of long-term vision, technocratic path dependencies, and a lack of political and public will (Brown and Farrelly, 2009). This appears to be a shared reality also in other places where current infrastructure planning is judged inflexible, narrow-minded, negligent of future uncertainties, as well as exclusive of broader goals, important stakeholders, and alternative paths of action (Ashley et al., 2008 (UK); Dominguez et al., 2009 and Störmer et al, 2009 (Switzerland); Economides, 2012 (USA); Ferguson et al., 2013 (Australia)).

In this thesis, alternatives that differ not only regarding their technology, but also regarding their governance and management, as well as spatial distribution will be considered. The water supply planning framework is embedded into an MCDA process (described below) to facilitate the consideration of multiple stakeholders and structure the decision including multiple, conflicting objectives (see chapters 2 and 6,Lienert et al., 2014b; Scholten et al., submitted).

Sustainable asset management

Infrastructure asset management approaches are gaining momentum in both water engineering research and practice (Cardoso et al., 2012; Christodoulou et al., 2008; Haffejee and Brent, 2008; Heather and Bridgeman, 2007; Marlow et al., 2010; Sægrov, 2005; Ugarelli et al., 2010; Vanier, 2001). Marlow et al. (2010) define asset management as *"A combination of management, financial, economical, engineering and other practices applied to (physical) assets with the objective of maximizing the value derived from an asset stock over a whole life cycle […]."* It is being applied at different spatial and temporal scales, be it strategic (long-term, 'planning for the future'), tactical (medium-term, 'determining which assets are to be replaced and how'), or operational (short-term, 'undertaking operations, maintenance, and operational risk management') (Marlow et al., 2010). The major reasons for the popularity as stated by 46 U.S. American utility managers are: a) provides better information about the age and condition of assets, b) helps to determine the level of maintenance needed to optimize asset performance, c) helps to assess the risks associated with the failure of various assets and to set priorities for their maintenance and replacement, as well as to d) understand the trade-offs and implications of management decisions, and e) use better information to justify proposed rate increases or capital investments (U.S. Government Accountability Office, 2004).

Even if it does considerably support decision making, asset management is no decision support framework in itself and needs to be supplemented by robust and feasible decision support tools (Alegre, 2010; Giustolisi et al., 2006; Selvakumar and Tafuri, 2012). These shall also facilitate the embedding of sustainability principles into business as usual asset management (Marlow et al., 2010). Marlow et al. (2010) found that *"strategic asset management, in conjunction with associated planning procedures, has the greatest scope for delivering against sustainability objectives"*, which an increasing number of water utilities is committing to. As regards the second decision problem, I therefore suggest to combine formal MCDA analyses with strategic infrastructure asset management approaches. In continuation of the works of Baur et al. (2003) and Carriço (2012) who used multi-criteria outranking methods for short to mid-term rehabilitation planning, a combination of strategic asset management with multi-attribute utility theory (MAUT) is presented in this thesis (Scholten et al., 2014; chapter 5). The results also form the basis for the selection of rehabilitation options within the 'main MCDA' (about good water supply infrastructure). By this, the complexity can be reduced and decisions-within-decisions (Gregory et al., 2012a) be avoided.

Managing uncertainty

Walker et al. (2003) provide one of the most comprehensive concepts for defining and managing uncertainty in model-based decision support. They define uncertainty as "*any deviation from the unachievable ideal of completely deterministic knowledge of the relevant system*". Three dimensions of uncertainty – location, level, and nature – are distinguished. The *location of uncertainty* refers to the step or part of the modeling process in which uncertainty arises, e.g. in the context, structure, and implementation of the model, in the input data, uncertainty of the parameters, or model outcomes. The *level of uncertainty* describes the degree to which something is uncertain, with a range spanning from statistical uncertainty over scenario uncertainty and recognized ignorance to total ignorance. Lastly, the *nature of uncertainty* describes the reason of the occurrence of uncertainty which can be either due to the imperfection of knowledge (*epistemic uncertainty*) or stochastic uncertainty caused by the inherent model variability (Walker et al., 2003). Stochastic uncertainty is sometimes also referred to as *aleatory uncertainty* (e.g. Kiureghian and Didevsen, 2009). In order to deal with these uncertainties, Walker et al. (2003) propose an uncertainty matrix for assessing the uncertainties in a systematic manner. The matrix' rows correspond to the location of uncertainty (e.g. the context, parameters etc.) which may have different levels and natures of uncertainty (arranged in columns). After this assessment, qualitative or quantitative uncertainty and sensitivity analyses can be performed to help allocate resources to the study of the most important uncertainties (Walker et al., 2003).

In the more quantitative modeling approaches, the aim of *uncertainty analysis* is to describe and quantify the (mostly statistical) uncertainty of the model output which results from the propagation of model parameter and input uncertainty (French, 2003). *Sensitivity analysis,* on the other hand, is used to determine how the output uncertainty is influenced by the uncertainty of the input or model parameters (Saltelli et al., 2004). Both uncertainty and sensitivity analysis use probabilities to describe uncertainty, which is propagated to the model outcomes. These approaches have been extended by *scenario planning*, which aims at quantifying the consequences of scenaric uncertainties (Schnaars, 1987; Schoemaker, 1995). Bradfield et al. (2005) identified three main schools of scenario planning that evolved over the past three decades: 1) intuitive logistics, 2) La Prospective, 3) probabilistic modified trends (PMT). All three schools require scenarios to be coherent, plausible, internally consistent, and logical. Their most important difference is the assignment of probabilities to the scenarios, besides minor dissimilarities in the scenario generation and evaluation process. The La Prospective and PMT methods assign probabilities to the scenarios, e.g. a base case and an upper and lower limit scenario (Bradfield et al., 2005). In contrast, scenarios in the sense of the intuitive logistics school are assumed equally probable or simply as possible visions of the future without trying to predict it (e.g. Ringland, 2002; Schnaars, 1987; Schoemaker, 1995).

When experts or decision makers provide inputs to the modeling and decision support process, the elicitation of uncertain knowledge (or uncertain preferences) remains a challenge (O'Hagan and Oakley, 2004), for more recent reviews see Krueger et al. (2012), Kynn (2008), and Low Choy et al. (2009). This is difficult not only due to the trouble of making precise probabilistic statements per se (O'Hagan et al., 2006), but also because the presentation of uncertainty affects judgment, leading to heuristic biases which contradict the mathematical laws of probability or other stipulations of rationality underlying decision support models (e.g. as first studied by nobel prize winners Tversky and Kahneman, 1974; 1981). In the decision sciences, different theories for the consideration of decision makers' attitude towards risk (known cause-effect, probabilistically quantifiable) and uncertainty (known cause-effect, not probabilistically quantifiable) have been proposed (e.g. see reviews in Abdellaoui et al., 2007; Bleichrodt et al., 2001; Cox et al., 2012; Eisenführ et al., 2010).

As the term 'uncertainty' itself already leads to considerable misunderstanding even within disciplines (Walker et al., 2003), I make no distinction between uncertainties elsewhere referred to as risk, uncertainty, ignorance, aleatory or epistemic uncertainty. Instead, I use the term uncertainty as in common language, i.e. to describe a situation where "knowledge gaps or ambiguities […] affect our ability to understand the consequences of decisions" (Gregory et al., p. 127). Probabilistic modeling is used throughout this thesis to quantify and propagate uncertainty of the presented decision-support models. Moreover, the uncertainty of changing future framework conditions is considered by including scenario planning following the intuitive logistics school (e.g. Schnaars, 1987; Schoemaker, 1995), which is increasingly common in decision analysis as well as water supply planning and infrastructure management (e.g.Cardoso et al., 2012; Goodwin and Wright, 2001; Montibeller et al., 2006; Stewart et al., 2013). Local and global sensitivity analyses are used to analyse the stability of model results to changes of the model parameters and to identify most influential parameters.

Multi-criteria decision analysis

To address the above research questions, multi-criteria decision analysis (MCDA) is proposed to complement the current GWP for long-term infrastructure planning because of its reported capability to (a) provide transparency and accountability, (b) resolve conflicts and integrate multiple perspectives, (c) enhance multi-stakeholder engagement and community participation, and (d) provide rich information, based on the (e) use of formal axioms to inform choices which ensure a logical and robust analysis (Hajkowicz and Collins, 2007). Sometimes confounded in the applied literature, MCDA is not just a single method, but an established research discipline originating from earlier works in economics, decision analysis, and operational research (Köksalan et al., 2013). It covers a range of different formal approaches which "*seek to take explicit account of multiple, conflicting criteria, help to structure the decision problem, provide a model that can serve as a focus for discussion and offer a process that leads to rational, justifiable, and explainable decisions"*(Belton and Stewart, 2002). Informative overviews of these methods are given in Belton and Stewart (2002) and Figueira et al. (2005).

The application of MCDA in water research and other fields has significantly increased over the past decade (Hajkowicz and Collins, 2007; Huang et al., 2011; Wallenius et al., 2008). Most applications in the water field concern water resources management and policy (e.g. Calizaya et al., 2010; Duckstein et al., 1994; Hajkowicz and Collins, 2007; Langhans et al., 2013; Marttunen and Hämäläinen, 2008; Raju and Pillai, 1999; Reichert et al., 2007). Less common is the application of MCDA in urban water (Afify, 2010; Chen et al., 2012; Joubert et al., 2003; Kodikara et al., 2010; Lindhe et al., 2013; Sa-nguanduan and Nititvattananon, 2010) and wastewater management (Ana et al., 2009; Borsuk et al., 2008; Ganoulis, 2003; Keeney et al., 1996; Lienert et al., 2011). There are at least two experiences reported from water and wastewater infrastructure rehabilitation planning (Baur et al., 2003; Carriço et al., 2012); both of which chose methods from the 'outranking school' (see Roy, 1991) of MCDA methods.

In a situation where infrastructure decisions are reportedly perpetuating the status quo and potential alternative options are ignored throughout the process (Ashley et al., 2008; Störmer et al., 2009), it seems self-evident to choose an approach which is value-focused (Keeney, 1992). Following a value-focused approach implies that alternatives are compared based on preferences about the importance and degree of achievement of objectives, i.e. what someone *really* wants to achieve by means of choosing an option, instead of comparing and ranking different options directly. For example, the objective 'high drinking water quality' can be achieved by various options such as advanced water treatment in a centralized facility or in-house water purification systems, switching the raw water source, or buying bottled water in the supermarket. Value-focused thinking is fulfilled by multi-attribute value and / or utility theory (MAVT / MAUT; where MAUT is used if the situation involves risky choices). Other desirable properties which MAVT / MAUT satisfy are (for details see Schuwirth et al., 2012a): 1) foundation on axioms of rational choice, 2) explicit handling of prediction uncertainty and stakeholder risk attitudes, 3) ability to process many alternatives without increased elicitation effort, and 4) possibility to include new alternatives at any stage of the decision process.

Keeney (1982) proposed a step-wise procedure for guiding the decision process:

1. Structure the decision problem
2. Assess possible impacts of each alternative
3. Determine preferences of decision makers
4. Evaluate and compare alternatives

Step 1 can be further subdivided into (see also Eisenführ et al., 2010; Gregory et al., 2012a; Keeney and Raiffa, 1993): 1a) structuring the decision problem (framing of the decision problem and setting of boundary conditions, defining who should be involved), and 1b) setting up the objectives hierarchy consisting of objectives and measurable attributes, and defining the decision alternatives. A thorough dealing with this task is crucial for the overall MCDA process since it shapes and influences the way the decision is made (Belton and Stewart, 2002; Morton and Fasolo, 2009). Close stakeholder interaction is required in this step, and a range of approaches exist to support individuals and groups in this endeavor (Belton and Stewart, 2010; Franco and Montibeller, 2011; Gregory et al., 2012a). The structuring of the case study underlying this thesis (see below) was conducted in form of face-to-face interviews and stakeholder workshops (described in detail in Lienert et al., 2014b; chapter 2 of this thesis). The definition of attributes was supported by additional expert interviews (Scholten et al., submitted; chapter 6). The selection of the most important stakeholders was based on a stakeholder and social network analysis, described in Lienert et al. (2013).

In **step 2**, the performance of decision alternatives regarding different objectives is assessed. Depending on the attribute to be quantified, this is accomplished either by means of detailed models (e.g. rehabilitation demand, groundwater recharge), expert estimation (e.g. water quality), or simple forward calculations (e.g. area demand). The assessment of asset performance and rehabilitation demand requires detailed failure and rehabilitation models, many of which are of probabilistic nature (see reviews in Kleiner et al., 2009; Kleiner and Rajani, 2001; Liu et al., 2012). These models have been successfully applied in larger networks, e.g. (Alvisi and Franchini, 2010; Eisenbeis et al., 1999), but are infeasible in most Swiss networks due to their high data demand for calibration (i.e. necessary sample size). In addition, these models do not consider common data particularities which may result in biased predictions. These particularities are left-truncation and right censoring of data due to incomplete recording over part of the assets' lifetimes, and selective survival bias, caused by the deletion of failure and replacement information of replaced assets (Le Gat, 2009; Mailhot et al., 2000; Renaud et al., 2012; Scheidegger et al., 2011). These difficulties are overcome by 1) the use of Bayesian instead of Frequentist parameter estimation to overcome the problems of small sample size, as done by (Dridi et al., 2009; Watson et al., 2004), and 2) the development of a failure model which is able to deal with left-truncated, right-censored, and selective survival data (Scheidegger et al., 2013; chapter 4). One major difficulty in Bayesian inference is the specification of sensible priors (Berger, 1990; Bousquet, 2008). Prior distributions for the models used in this thesis are obtained in two ways: firstly by quantitative elicitation from experts ('expert elicitation'; Scholten et al., 2013; chapter 3), and secondly by inference from three larger water utilities (Scholten et al., 2014; chapter 5). Expert elicitation has been widely used to inform model parameters in environmental and engineering modeling (Krueger et al., 2012; Kynn, 2008; Low Choy et al., 2009) and numerous guidelines exist (Lele and Allen, 2006; O'Hagan et al., 2006). It seems however, that the explicit consideration of imprecision in elicitation has been overlooked in the past. Extending the approach of Oakley and O'Hagan (2010), and O'Hagan et al. (2006) to the elicitation of imprecise quantiles of cumulative probability distributions, pipe survival distributions are obtained from eight experts and aggregated to obtain an intersubjective prior probability distribution of the group of experts. This is then used for Bayesian parameter inference of the pipe survival model parameters. Instead of expert priors, the failure model is parameterized with prior distributions from larger water suppliers and then combined with a rehabilitation model to evaluate the performance of different rehabilitation alternatives. A number of failure and rehabilitation models exist, e.g. KANEW (Kropp and Baur, 2005), PiReM (Fuchs-Hanusch et al., 2008), D-WARP (Kleiner and Rajani, 2004), Aware-P (Cardoso et al., 2012), Casses (Renaud et al., 2012), WilCO (Engelhardt et al., 2003), or PARMS Planning (Burn et al., 2003). Because none of these models meets three core requirements of the approach herein presented, namely combinability with the newly developed failure model, flexible implementation of rehabilitation strategies, and propagation of parameter uncertainty to the results, a sector-independent asset management model (FAST, Fichtner Asset Services & Technologies, 2013) is used.

The elicitation of stakeholder preferences (**step 3**) for preference modeling in MAUT aims at quantifying three components: a) the relative importance of objectives, which is estimated by assigning weights to them, b) the marginal value functions, which convert attribute levels (e.g. costs in CHF/year) to a neutral scale between 0 and 1, and c) an overall aggregation function to combine the predictions of all attributes (e.g. costs and water quality) to an overall judgment of 'good water supply infrastructure'. Because the elicitation of preferences is cognitively very demanding and time-intensive, often only some of these components are obtained. This can for instance be supported by interactive online software (Marttunen and Hämäläinen, 2008; Mustajoki and Hämäläinen, 2000), questionnaires and surveys (Hyde et al., 2011), and individual or group interviews (Bleichrodt et al., 2001; Karvetski et al., 2009a; Smidts, 1997). Nearly all practical MCDA interventions entail the elicitation of relative importance weights. Sometimes, the value and / or utility functions are elicited in detail or otherwise assumed as linear (Raju and Vasan, 2007; Weber, 1987). The aggregation functions (or rather their underlying preferential dependencies) are only rarely elicited (examples are Keeney and Raiffa, 1993; Raju and Pillai, 1999); a linear aggregation function is commonly used instead (Hajkowicz, 2008; Hyde et al., 2005; Joubert et al., 2003). Research has shown that elicitation is prone to behavioral or cognitive biases, their severity depending on the method used (Bleichrodt et al., 2001; Borcherding et al., 1991; Morton and Fasolo, 2009; Weber and Borcherding, 1993). As regards for example weight elicitation, comparisons of methods revealed important inconsistencies, but no definite conclusion regarding the best method emerges from the literature (Borcherding et al., 1991; Weber and Borcherding, 1993). To address this, Borcherding et al. (1991) propose to perform consistency checks using different methods while others recommend asking for imprecise instead of precise estimates (Jessop, 2011; Mustajoki et al., 2006). The descriptive fallacy of utility theory in addressing biases linked to subjective probabilities (e.g. Morgan and Henrion, 1990; Spetzler and Stael von Holstein, 1975), preferences about risky choices (Allais paradox; Allais, 1953), and choices under ambiguity (Ellsberg paradox; Ellsberg, 1961) have led some to develop alternative theories for decisions under risk (e.g. Abdellaoui et al., 2008; Chateauneuf and Cohen, 2000; Schmidt et al., 2008) if not finding ex-post remedies (Bleichrodt et al., 2001). Yet, others argue that the descriptive fallacy does not disqualify the use of utility theory in prescriptive analyses (e.g. Belton and Stewart, 2002; French, 2003). In this work, I follow the later rationale, but aim at increasing robustness through performing consistency checks, and by allowing for imprecision in preference statements. Due to the large number of preference components to be elicited, some propositions for simplified elicitation are made (Scholten et al., submitted; chapter 6). Given the restricted time available with the stakeholders, a part of the interview was focused on specific, highest-ranked objectives only, as identified during a preceding online survey. For the main MCDA, preferences of ten stakeholders from different professional backgrounds (e.g. operating staff, infrastructure engineering, administration, water chemistry) and decision-making levels (local, regional, cantonal and national), were assessed. The smaller MCDA (about finding good strategic rehabilitation alternatives, using a reduced set of three fundamental objectives) is done using simple preference assumptions and the robustness of the resulting ranking against changes of the supposed preferences is verified applying local sensitivity analyses (Scholten et al., 2014; chapter 5).

Lastly, in **step 4**, the alternatives are compared based on the attribute outcomes of each alternative (e.g. how much it costs, how reliably the water is supplied) and on individual stakeholder preferences (e.g. how important are 'low costs' compared to 'good water supply'). Uncertainty analysis and simulation studies are common means to quantify model output uncertainty (of the ranking of the alternatives) through propagation of model parameter input uncertainty (French, 2003), also in MCDA studies. Local sensitivity analyses (LSA) is commonly used to assess the stability of results given changes in inputs and / or parameters. Global sensitivity analyses (GSA) are used to study how output uncertainty can be apportioned to different sources of uncertainty in the model input (Saltelli et al., 2004), but are much less common in MCDA studies (Gómez Delgado and Bosque Sendra, 2004; Saltelli et al., 1999a). Most of the available studies, however, are limited in breadth, considering the uncertainty of the weights alone (Butler et al., 1997; Hyde et al., 2005; Jessop, 2011; Jiménez et al., 2006; Mustajoki, 2012; Mustajoki et al., 2006; Raju and Pillai, 1999), or weights in combination with attributes (Gómez Delgado and Bosque Sendra, 2004; Hyde et al., 2004; Saltelli et al., 1999a). A more comprehensive analysis of uncertainty arising from uncertain preference parameters regarding the attributes, weights, value and utility function curvature, as well as aggregation functions, has not been reported so far and is done for the main MCDA in this thesis (Scholten et al., submitted; chapter 6). To additionally capture context uncertainties, such as population growth, urbanization, and socio-economic change, scenario planning (Schnaars, 1987; Schoemaker, 1995) is combined with MCDA. This combination is increasingly used in MCDA interventions (e.g. see review in: Stewart et al., 2013). Within this thesis, the evaluation is done under four future scenarios following the approach suggested by Goodwin and Wright (2001), and the robustness of the resulting rankings is discussed.

The overall approach is developed and validated in a case study in a rural area near Zurich, Switzerland: (the 'Mönchaltorfer Aa'; descriptions follow in chapter 2, chapter 5 (one water utility only), and chapter 6).

1.4 Thesis outline

The decision problem of the main MCDA and its structuring are described in **chapter 2** (Lienert et al., 2014b; under review). Here, also a description of the structuring process, the alternatives, objectives, and attributes is presented. It furthermore presents the case study and provides insights into the creation and details of the four future scenarios used during later MCDA preference modeling and evaluations

Chapters 3 and 4 address the shortcomings in the prediction of water main deterioration, pipe failures, and expected service life of pipe assets under the common data situation in Switzerland and, more specifically, small water networks. Therefore, in **chapter 3** (Scholten et al., 2013; published), a simple pipe survival model is proposed. Priors for Bayesian parameter estimation of pipe survival are quantitatively inferred from eight experts, and aggregated into one distribution under consideration of within and in-between experts uncertainties. These are then combined with different amounts of pipe survival data to explore the effect of sample size and demonstrate the applicability of the approach in data-scarce situations where Frequentist parameter estimation is unsatisfactory. Because replacement data alone do not allow to differentiate between managerial and structural reasons for replacement, and are furthermore seldomly available, this is followed by the presentation of a new modeling framework in **chapter 4** (Scheidegger et al., 2013; published). This framework explicitly addresses the absence of failure data from replaced pipe and hence reduces this selective survival bias in pipe failure and service life prediction. Additionally, a simple pipe failure model intended for use in small water networks is presented.

In **chapter 5** (Scholten et al., 2014; published) an approach to compare long-term rehabilitation strategies based on a combination of failure and rehabilitation modeling for strategic asset management with MAUT is developed and presented. The modeling groundwork from chapter 3 and chapter 4 builds the basis for rehabilitation modeling and performance predictions regarding three criteria (costs, system reliability, intergenerational equity). Local sensitivity analyses are done to assess the robustness of the ranking of the 18 rehabilitation alternatives. The approach is exemplified for a small water supplier of the "Mönchaltorfer Aa" case study and specific results are presented.

Finally, different sources of uncertainty arising in MAUT analyses and an approach how to deal with them in practice are presented in **chapter 6** (Scholten et al., submitted). The decision problem as presented and structured in chapter 2 is used to exemplify the elicitation and modeling of preferences of ten stakeholders. Based on these preferences for 30 objectives, and predictions for the corresponding attributes, 11 water supply alternatives are compared under four future scenarios.

The conclusions of the presented work are drawn in **chapter 7**. The strengths and weaknesses of the approach are put into the context of water infrastructure planning in Switzerland and other countries. Finally, further potential for future research is identified.

2 Structured decision making for sustainable water infrastructure planning under four future scenarios

Judit Lienert, Lisa Scholten, Christoph Egger, and Max Maurer

Swiss Federal Institute of Aquatic Science and Technology, Eawag, P.O. Box 611, CH-8600 Duebendorf, Switzerland

EURO Journal on Decision Processes (under review)

Abstract

Water supply and wastewater infrastructures are vital for human well-being and environmental protection; they adhere to highest standards, are expensive and long-lived. Because they are also aging, substantial planning is required. Climate and socio-economic change creates large planning uncertainty and simple projections of past developments are no longer adequate. This paper aims to guide through the initial phases of a Structured Decision Making (SDM) procedure to increase the sustainability of water infrastructure planning. Our SDM-procedure includes various stakeholders in an exemplary Swiss case study. We evaluate the SDM-approach critically based on stakeholder feedback, give general recommendations and provide ample material to make it applicable to other settings. We carried out 27 interviews and two stakeholder workshops. We identified important objectives for water infrastructure planning, including all three sustainability pillars and their respective attributes (indicators, benchmarks) to measure how well objectives are achieved. We then created strategic decision alternatives, including "business-as-usual" upgrades of the central water supply and wastewater system, but also semi- to fully decentralized alternatives. To tackle future uncertainty, we developed four socio-demographic scenarios. We use these to test the robustness of decision alternatives in a later MAUT-analysis. Additionally, we contribute to the topical discussion of combining scenario planning with MCDA and demonstrate how scenarios can stimulate creativity when generating decision alternatives. Their internal consistency is ensured by rigorously specifying them with help of a strategy generation table. Our SDM-procedure can be adapted to inform decisions about sustainable water infrastructures in other contexts.

Keywords

Decision making, water infrastructure, scenario planning, stakeholder participation, structuring, water management

2.1 Introduction

2.1.1 Structured decision making

Decision making in environmental management is complex. Typically, this affects various actors and requires difficult trade-offs across a wide range of environmental and socio-economic objectives. If future generations are affected, long time spans need to be considered. Multi-Criteria Decision Analysis (MCDA) provides a useful framework for making better-informed, more sustainable and participatory decisions (e.g. Clemen and Reilly, 2001; Eisenführ et al., 2010; Keeney, 1992; Keeney and Raiffa, 1976). There are numerous examples of environmental applications (reviewed in Huang et al., 2011; Linkov and Moberg, 2012).

To support the choice between decision alternatives, mathematical models can be applied that integrate the decision makers' (subjective) preferences for outcomes with the (objective) performance of the alternatives on a set of previously determined objectives. However, as nicely outlined in a book by Gregory et al. (2012a), it often suffices to structure the decision together with the stakeholders to clarify the trade-offs and find an agreement between parties. This structuring process then may – or may not – be followed by a formal MCDA, whereby modeling and expert knowledge to predict outcomes is combined with stakeholder preferences.

In this paper, we focus on the initial Structured Decision Making steps (SDM; Gregory et al., 2012a) that are crucial in any decision, but are often neglected. Usually, following steps are carried out (see textbooks, e.g. Clemen and Reilly, 2001; Eisenführ et al., 2010; Gregory et al., 2012a; Keeney, 1992; Keeney and Raiffa, 1976): (1) clarify the decision context; (2) define objectives and attributes; (3) develop alternatives; (4) estimate consequences; (5) evaluate trade-offs and select alternatives (this is a combination of the decision makers' subjective preferences with the objective consequences of the alternatives); and (6) implement, monitor and review. Gregory et al. (2012a) argue that good decision making does not always require quantitative modeling, but that structuring the decision consistent with sound theory helps to discipline thinking and to make decisions more transparent. Often, a linear additive value model is used to calculate an overall value for each alternative (step 5), based on a weighted sum of the alternatives' consequences for each attribute (following multi-attribute value or utility theory, e.g. Eisenführ et al., 2010; Gregory et al., 2012a; Keeney and Raiffa, 1976). All models require attributes (indicators) that make objectives measurable, a prediction to quantify how well each alternative fulfills the objectives, and preference information from decision makers. Hereby, each attribute receives an importance weight and a value function transforms attribute levels to a neutral scale between 0 and 1. Alternatives that achieve the highest values[1] are proposed and discussed with the decision makers.

2.1.2 Combining scenario planning with MCDA

Water infrastructures are long-lived, with average pipe lifespans of water supply and sewer pipes of e.g. 80 years (Martin, 2009). It is thus especially important to consider intergenerational equity, which is a core aspect of sustainable development[2] (WCED, 1987; for a conceptual discussion see Wuelser et al., 2012). For such time ranges, the future is "deeply uncertain"[3] and it is impossible to use probabilistic models (e.g. Walker et al., 2010). However, most MCDA-methods are deterministic, and the uncertainties are often

---

[1] Formally, the linear additive value model is: $$v(a) = \sum_{i=1}^{m} w_i v_i (a_i)$$

where:	v(a)	=	total value of alternative a
	a_i	=	attribute level of alternative a for attribute i
	$v_i(a_i)$	=	value for attribute i of alternative a
	w_i	=	weights (or scaling constants of attribute i, and sum of w_i equals 1

[2] "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs." (WCED 1987; p. 43)

[3] "Level 3 uncertainty represents deep uncertainty about the mechanisms and functional relationships being studied. We know neither the functional relationships nor the statistical properties, and there is little scientific basis for placing believable probabilities on scenarios. In the case of uncertainty about the future, level 3 uncertainty is often captured in the form of a wide range of plausible scenarios. Level 4 uncertainty implies the deepest level of recognized uncertainty; in this case, we only know that we do not know." (Walker et al. 2010; p. 918).

internal (epistemic uncertainty or imprecision; reviewed in Stewart et al. 2012; also see Reichert et al 2014)[4].

Scenario building is a tool to systematically explore the future, without trying to predict it (e.g. Ringland, 2002; Schnaars, 1987; Schoemaker, 1995). Early examples come from business strategy formation (e.g. the famous Shell example: Wack, 1985). There are also numerous environmental applications (e.g. Peterson et al., 2003; Swart et al., 2004), including strategic planning for urban water infrastructures (Dominguez et al., 2011; review in: ; Lienert et al., 2006; Störmer et al., 2009; Truffer et al., 2010; review in: Dong et al. 2012).

Recently, researchers started combining scenario planning with MCDA. The combination is not trivial, because it adds an additional dimension to the already highly complex MCDA-analyses. One problem is how to include stakeholder preferences. If it is assumed that the preferences differ under different scenarios, a value function for each decision maker under each scenario must be constructed (Karvetski et al., 2009a; 2011; Lambert et al., 2012; Montibeller et al., 2006; Stewart et al., 2013), so that the elicitation process becomes more laborious (Ram and Montibeller, 2012; Ram et al., 2011). As shortcut, only shifts in the relative importance of certain value function components, compared to a baseline value function can be elicited (e.g. Karvetski et al., 2009a; 2011; Lambert et al., 2012). Stewart et al. (2013) propose to aggregate across scenarios by introducing "metacriteria", but this approach remains to be tested in practice.

2.1.3 Water infrastructure planning

The water supply and wastewater disposal infrastructures are crucial for the provision of clean water and water for firefighting, urban hygiene, protection against flooding, and water pollution control. In many OECD countries, the infrastructures adhere to highest standards and are expensive; the replacement values of the public wastewater system (excluding household connections) are typically between 2'600 and 4'800 US$ per person (Maurer et al., 2005). The annual investment need in OECD countries in the water sector is approximately 0.75 % of the GDP (Cashman and Ashley, 2008), which translates into 300'000 million US$ annually (OECD 2012). Despite its success in the industrialized world, the centralized infrastructure system is increasingly criticized for its lack of sustainability (e.g. using clean water to flush toilets, loss of nutrients, e.g. phosphate that could be recycled). The central system with extensive underground pipe networks and large treatment plants is also very inflexible. Decentralized options for water supply and wastewater disposal are gaining increasing momentum in the engineering community (e.g. Guest et al., 2009; Larsen et al., 2009; 2012; Libralato et al., 2012a).

Despite long service lives, infrastructures are often planned with mid-term projections (<25 years) of past developments. This approach is deficient in that it does not account for future developments. Under climate change, we expect severe droughts and more-frequent heavy storms in Central Europe (e.g. Kysely et al., 2011). For example, the sewers may have increasing difficulties to reliably drain storm water, resulting in more combined sewer overflows that pollute rivers and lakes, and more urban floods (e.g.

4 Stewart et al. (2013; pp. 683–684) distinguish "internal uncertainty" from "external driving forces". Internal uncertainties concern e.g. the imprecision of measurements; probability frameworks can deal with these. Stewart et al. (2013) also classify epistemic uncertainty as internal uncertainty. In epistemic interpretations, probabilities can be used to quantify human (expert) knowledge or belief concerning the probability of something occurring. How to conceptually deal with uncertainties in environmental management with a specific focus on MCDA is discussed by Reichert et al. (2014). In contrast, external uncertainties may much more strongly affect the outcome of decisions we make today. These uncertainties (e.g. future climate, demographic or economic development) can often be better captured by the scenario approach.

Arnbjerg-Nielsen K and H.S., 2009; Butler et al., 2007; Patz et al., 2008). Socio-demographic and economic pressure add to planning uncertainty – "the challenge is daunting" (Milly et al., 2008).

We know of only few applications of MCDA in urban water infrastructure planning for OECD countries. Most MCDA-projects in the water sector concern water policy and water resources management (e.g. Hämälainen et al., 2001; Reichert et al., 2007). Also for infrastructures, mainly water resources management, including hydroelectric power schemes are considered (e.g. Eder et al., 1997; Kodikara et al., 2010), but rarely urban drinking and wastewater management (see review by Hajkowicz and Collins, 2007 and an early example by Keeney et al. 1996). From the water engineering sector, there is growing interest in comparing different infrastructure options using "indicators", usually with Life Cycle Analysis (LCA) (Balkema et al., 2001; Lundie et al., 2004), sometimes in combination with Multi-criteria analysis (MCA; Palme et al., 2005). The indicators cover environmental and social criteria such as "acceptance" (of phosphorus products from sewage), "reliability of service" and "working conditions" (Palme et al., 2005). In one case, non-conventional decentralized options were evaluated, but based on purely environmental indicators (Lundie et al., 2004). However, sustainability indicators remain an "elusive concept" (Ashley et al., 2008). To our knowledge, developing a comprehensive objectives hierarchy, based on multi-attribute value theory (MAVT; e.g. Eisenführ et al. 2010; Keeney 1992; Keeney and Raiffa 1976) for use in a full MCDA-analysis that accounts for long-term changes is new in the field.

2.1.4 Objectives of this paper

The aim is to present and critically discuss the initial SDM-structuring and decision-making steps, based on Gregory et al.(2012a) and on stakeholder feedback. As illustration, we use a complex real example of water infrastructure planning in Switzerland. We include a broad range of stakeholders and develop a comprehensive set of decision objectives, diverse alternatives and four future scenarios. Although developed in a local stakeholder process, we set up the SDM-framework so that it can be adapted to water infrastructure decisions in other countries.

2.2 Materials and Methods

2.2.1 Step (1) Clarify decision context

In the first step of the SDM-process, the decision context, scope and boundaries of the decision problem are clarified. Hereby, "decision sketching" may be useful, as illustrated with examples from environmental management by Gregory et al. (2012a). For structuring, e.g. means ends networks, preliminary objectives hierarchies, consequence tables, influence diagrams or decision trees are suggested (also see Clemen and Reilly, 2001; Eisenführ et al., 2010). In this step, one must also decide who should participate.

The here presented project was clearly centered on scientific research rather than on finding solutions for a single decision problem. We aimed at providing a procedural tool for "Sustainable Water Infrastructure Planning" (SWIP, 2013) that enhances planning efficiency, can deal with uncertainty and is well accepted.

Case study

To mirror the research with real stakeholders, we identified a suitable case study, structured the case study project including different types of stakeholder participation, and identified the stakeholders that should be involved. We carried out our research in a region "Mönchaltorfer Aa" near Zurich with four smaller communities and about 24'200 inhabitants. There is extensive agriculture, but also urban growth pressure from Zurich. The nearby Lake Greifensee is an important recreational and nature protection area. It is one of the few Swiss lakes that still have phosphate eutrophication problems, stemming from wastewater and agriculture (AWEL, 2003, 2006). In summer, there is danger of fish kills due to oxygen depletion in deeper water layers and high temperature in surface layers (AWEL, 2003).The discharge from wastewater treatment plants (WWTPs) into smaller rivers upstream of Lake Greifensee results in inadequate river water quality and contains micropollutants (AWEL, 2006).

Stakeholder selection

The SDM-process is intended for groups of five to twenty-five people that work intensively on a complex problem (Gregory et al., 2012a). A decision sketch can also help identifying stakeholders. After clarifying which environmental and societal endpoints are affected by the decision alternatives, one can ask: "who will care about these outcomes?" However, Gregory et al. (2012a) provide little guidance on methods to choose these participants.

In our application, we placed much more emphasis on selecting stakeholders than usually reported in the SDM-literature. To identify those who play a role in water infrastructure planning or who could be affected, we carried out a stakeholder and social network analysis (Lienert et al., 2013). We found that over 40 actors play a role, with a clear dominance of local and engineering actors. The network analysis confirmed the hypothesis of a strongly fragmented water sector, namely between water supply and wastewater (and others), and between decisional levels[5]. We used this work to select the workshop participants and interview partners in the here presented paper. Besides obvious stakeholders such as the local planning engineers and municipalities, also representatives that were perceived to be less important were included, e.g. from the cantonal and national authorities (details see Lienert et al. 2013a).

2.2.2 Step (2) Define objectives and attributes

General procedure to create the objectives hierarchy

Objectives define "what matters" in the decision and attributes (performance measures / indicators) make objectives operational (Gregory et al., 2012a). Objectives can be organized hierarchically and provide a framework to transparently compare the performance of alternatives. It is crucial that the decision makers (in our example the selected national, cantonal, and national stakeholders) understand and accept objectives and attributes; and also that specific rules are followed: objectives should comprehensively cover the decision, be fundamental, concise and sensitive, meaning that they should help distinguish between alternatives (e.g. if costs are the same in all alternatives, "low costs" are not suitably sensitive). They should be understandable, simple, non-ambiguous, non-redundant and preferentially independent (for the additive model; e.g. Eisenführ et al., 2010; Gregory et al., 2012a; Keeney and Raiffa, 1976).

Although people usually have a good idea about what is important to them, generating good objectives for environmental decisions is not trivial. Creativity techniques can help such as brainstorming a wish list, considering shortcomings or new perspectives (Clemen and Reilly, 2001). Based on environmental case study examples, Gregory et al. (2012a) recommend five steps (also see Keeney, 1992; Keeney and Raiffa, 1976): (1) brainstorm, (2) separate means from ends, (3) separate "process" or "strategic" from "fundamental" objectives, (4) build hierarchy and (5) test usefulness of objectives. It is crucial to avoid "means" objectives, which are important only to achieve a more fundamental objective. Hereby, means-ends networks can be useful (nicely illustrated inClemen and Reilly, 2001; Gregory et al., 2012a). If much is known, a top-down creation of the objectives hierarchy is recommended; it helps to ask: "what do you

---

5 These are e.g. local practitioners (engineers or operating staff of treatment plants), representatives from administration and politics from municipalities, the region (e.g. cantonal agency for waste, water, energy and air) and the national level (e.g. environmental protection agency; associations of water professionals).

mean by that?" for a more-detailed description of an objective (Clemen and Reilly, 2001). If little is known, it is suggested to move from lower to higher levels of the hierarchy.

To quantify objectives, attributes are needed (Eisenführ et al., 2010; Gregory et al., 2012a). "Natural" (e.g. $, hours) are clearly preferable over "proxy" attributes, which operationalize objectives only indirectly. However, the latter often cannot be avoided in environmental management (e.g. using "area" to measure "species abundance"). Constructed attributes such as seven-point Likert scales (Likert, 1932), known from psychological questionnaires, may be useful in environmental decisions, too (Gregory et al., 2012a). However, expert judgments are rarely unambiguous. It is thus recommended to combine numerical scales with narrative descriptions ("defined impact scales").

Procedure as applied in the SWIP application example

In our application, the objectives hierarchy was generated in a multi-step, iterative procedure. This comprised a desktop analysis to create a preliminary objectives hierarchy (top-down approach), face-to-face interviews with stakeholders, and a stakeholder workshop. Our aim was to generate a generic hierarchy applicable also to other cases of water infrastructure planning.

A preliminary objectives hierarchy was set up by the project team, based on engineering requirements. 'Good water supply' includes the uninterrupted provision of drinking water in high quality and quantity, and water for firefighting. Objectives of the wastewater system include 'urban hygiene' and the 'protection of water bodies' as stipulated in environmental laws. We included 'low costs' and 'intergenerational equity' to cover all pillars of sustainability. We provide more details in Lienert et al. (2014a; see Supporting information part A [SI-A]).

Face-to-face interviews

We discussed these objectives in the 27 face-to-face interviews for the stakeholder and network analysis (Lienert et al., 2013). We described the purpose of objectives (help to choose among ten infrastructure decision alternatives) and their properties (see above). First, the interviewees freely stated which objectives they found appropriate, before showing our own highest-level objectives. We assigned their objectives to ours and asked whether they agreed or if a new top-level objective was required. In this way we worked through all branches of the hierarchy. To select objectives, we asked for an importance classification[6]. We also asked for ideas about attributes and for general feedback[7]. We categorized answers and calculated the number of comments in each category. The hierarchy was later thoroughly revised by the project team.

Stakeholder workshops

In our application example, we carried out two stakeholder workshops in the study region in April and May 2011 (each five hrs.). The first was a scenario workshop (see below); in the second we created alternatives (see below) and discussed the objectives. At this second workshop 20 people participated, identified by the stakeholder analysis, including representatives from different municipalities, sectors, institutions and companies at the local, cantonal and national level. We presented the objectives hierarchy and requirements for "good" objectives. This was familiar to most, thanks to the previous interview. Participants systematically worked through the hierarchy and discussed in pairs, which objectives they

---

6 Essential objectives (without this objective I cannot judge whether fundamental objective is reached), important (without this it is difficult…) and nice to have (attainment of fundamental objective can be judged without this).

7 Specific questions: "What would be next step and who should do it?" / "What are your expectations, fears or hopes w.r.t. our project and Eawag?" (Eawag is our research institute, i.e. the Swiss Federal Institute of Aquatic Science and Technology) / "Do you have general feedback, also concerning the interview or recommendations?"

found really fundamental or which were missing. We collected their notes and discussed the objectives in the plenum. Each participant was asked to designate points to those three objectives perceived as the least relevant. At the end of the workshop, we asked for feedback.

Final objectives hierarchy and attributes

After the workshop, the project team again revised the objectives and attributes. We decided for which attributes we could generate the prediction for each alternative ourselves (results of dimensioning and engineering models in SWIP, know-how, literature survey) and which required other expert information. For these, we asked one to four experts to define an adequate attribute, the worst- and best-possible values and the attribute levels of our decision alternatives (Tab. 1). If judgments differed strongly, we increased the ranges, namely for 'high co-determination of citizens' (two experts with different estimates). For 'flexible system adaptation', judged by four engineers, we calculated the average and standard deviation. Alternatives with more than 10 % deviation were discussed and the point of view defended (similar to a group Delphi; Schulz and Renn, 2009). Then a final score was assigned by the group, with larger interval ranges to depict higher uncertainty or variance.

2.2.3 Future scenarios

Creating future scenarios is not part of standard SDM-procedures. We introduced this step because a main aim of our project is to develop a decision procedure that can cope with uncertainty. We adapted four Swiss development scenarios from an earlier National Research Program (NRP 54; www.nfp54.ch) to our local case in the first stakeholder workshop in April 2011, following Truffer et al. (2010). We invited 22 members of the four communities, but not from the national and cantonal level, because we found that local people should adapt the scenarios to their specific case. The exclusion of high-ranking officers also helped to create a comfortable workshop feeling. The 15 participants came from all four communities; they represented both water sectors and different roles (i.e. political or technical-engineering focus). First, we presented the SWIP-study and scenario planning: to create a picture of the future that is internally consistent and plausible, but not necessarily desirable or probable; the scenario descriptions are based on key factors that may differ in each future world (Schnaars, 1987). The scenarios were depicted to the year 2050. They were discussed and adapted to the local case in three groups, in which we ensured an equal distribution of perspectives. Specifications were based on the variation of eight factors, relevant for water infrastructures. The scenarios were visualized, presented and discussed in the plenum (see Lienert et al., 2014a). Finally, participants gave feedback in the plenum concerning: "Which development would I be happy about?" and "What did I learn?"

2.2.4 Step (3) Develop alternatives

In simple decision problems, one often starts with defined alternatives; the SDM-procedure then aims at choosing the best, but environmental management situations are not usually simple. Often, alternatives are complex sets of actions that need to be created. The focus of SDM then "is all about the development of creative alternatives that are responsive to the defined objectives"(Gregory et al., 2012a).

Good alternatives should be complete, comparable and value focused (i.e. address key aspects), fully specified, internally coherent and distinct. Three basic steps are recommended: (1) brainstorm management responses, (2) organize these into fully specified alternatives and (3) refine iteratively (Gregory et al., 2012a). We also recommend Eisenführ et al. (2010), Keeney (1992) and especially Clemen and Reilly (2001)concerning creativity techniques. These include idea checklists, Osborn's 73 idea-spurring questions (Osborn, 1963), strategy generation tables (Howard, 1988), metaphorical thinking and many more.

"Morphological forced connection" is a creativity-technique where different factors characterizing a problem are brainstormed (e.g. financial strategy) and different specifications are listed under each factor (e.g. constant budget, progressive budget,…, Clemen and Reilly, 2001). Then, combinations and permutations are tried out. A "strategy generation table" (Howard, 1988) is a more stringent variant, where each decision alternative (=strategy) consists of exactly one chosen specification for each factor, which are combined. It is a good framework to easily screen all imaginable combinations for useful candidates. Examples come from business problems or a NASA space-exploration mission (see Clemen and Reilly, 2001 and references therein). Strategy generation tables seem especially suited for environmental management problems (e.g. Gregory et al., 2012a, b).

In our application example, we used a strategy generation table to create alternatives in the second stakeholder workshop in May 2011. We used the four socio-economic scenarios from the first workshop as background. Note that this was not necessary for the MCDA, since we will analyze the performance of all alternatives under all scenarios; it was just a way to stimulate creativity. We prepared the strategy generation table beforehand (see Lienert et al., 2014a). The 17 factors, which consisted of different specifications, concerned the organizational structure, geographic extent, financial strategy, construction and operation of infrastructure and system technology for wastewater and drinking water. The 20 participants were split into four mixed groups and assigned to a scenario. Each created at least two strategic alternatives by choosing a plausible specification for each factor. These backbones were used by the project team to develop detailed and internally consistent alternatives (feedback see 2.2.2 "stakeholder workshops").

2.3 Results

Below we present the results of the SDM-process as carried out in the Swiss case study. In the Discussion, we compare our approach with the more general SDM-procedure (Gregory et al., 2012a). The results of the stakeholder analysis are presented in Lienert et al. (2013) and some general aspects concerning this first structuring step in the Discussion below.

2.3.1 Step (2) Define objectives and attributes

Face-to-face interviews

In the stakeholder interviews, five of six fundamental objectives at the highest hierarchical level were perceived as essential or important by nearly all (see Lienert et al., 2014a). Some additional objectives were proposed. Most were already covered under a different title or were means objectives. For example, 'good state of infrastructure' is a means objective to achieve e.g. 'safe water supply and wastewater disposal'. Several suggestions included trade-offs that will be calculated in the MCDA (e.g. 'optimized cost-benefits'). We also considered 'transparency' to be covered by the SDM-procedure. We decided that 'protection of floodplains' is outside our systems boundary, but included the new objective 'professionalism' ('high quality of management and operations') in the revised hierarchy. We strongly discussed the objectives in the project team and developed a larger hierarchy.

Stakeholder workshop

The revised objectives hierarchy was discussed in the stakeholder workshop (see Lienert et al., 2014a). It was not possible to delete objectives; we had hoped that we could reduce the large hierarchy to a smaller, better manageable set. Objectives describing the classic infrastructure system ("safe drinking water supply", "safe wastewater disposal") were nearly undisputed. Most of the discussion focused on objectives characterizing decentralized water supply and wastewater treatment alternatives. For water supply, these were 'household water of good quality' (lower quality than drinking water for washing etc.; Fig. 1) and 'water for firefighting', which in Switzerland is combined with drinking water supply. For 'low costs', mainly the total annual costs were seen as important, but less 'low cost fluctuations' and 'easy fundraising', which we deleted later. Objectives of 'high social acceptance' and 'intergenerational equity' were most strongly questioned. Despite this, we kept most of the questioned objectives, because neither the plenary discussion nor the distribution of points provided a clear justification. Moreover, we found it important to include all pillars of sustainable development (Wuelser et al., 2012). We kept some objectives because they were necessary to distinguish between alternatives (e.g. 'flexible system adaptation' or 'low time demand for end users').

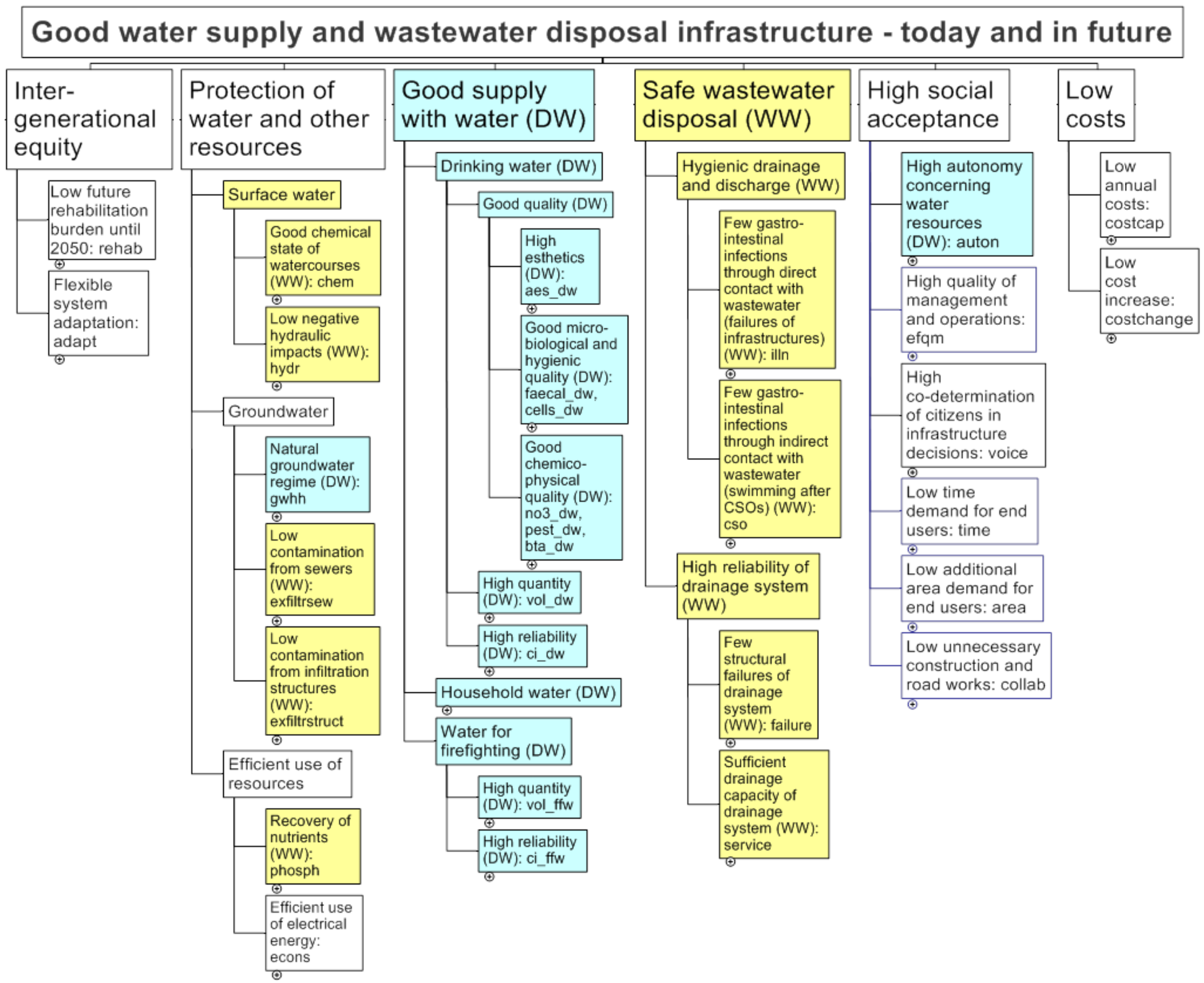


Figure 1: Final objectives hierarchy as used in the SWIP-project for water infrastructure decisions in the case study area Mönchaltorfer Aa. The objectives hierarchy is transferable to other cases. Objectives without shading are used for the entire network, objectives with blue shading only apply to the water supply infrastructures (DW = drinking water), and objectives with yellow shading only to the wastewater infrastructures (WW = wastewater). After the colon, the short name(s) of the respective attributes are given. CSOs = combined sewer overflows (discharge of mixed rain and wastewater to water bodies without / with only basic treatment).

Final objectives hierarchy and attributes

The fundamental objectives of the final hierarchy are given in Figure 1 and the attributes in Table 1 (details including ranges, narrative descriptions and Status Quo see Lienert et al. 2014). The objectives hierarchy and attributes were constructed to analyze the case study, but also to be applicable to other water infrastructure planning decisions; i.e. we consider it to be as exhaustive as sensibly possible. To make the work manageable in our SWIP-project, we split the water supply and wastewater system (three PhD students and one postdoc work on the project), but collaborated tightly to come up with a holistic hierarchy.

Table 1: Attribute description to measure how well an objective is achieved. For each fundamental objective (bold), we give the short name (see Fig. 1), units, and describe the corresponding attribute(s). Details concerning the calculations, ranges, and the Status Quo are given in Lienert et al. 2014. WWTP = wastewater treatment plant, DW = drinking water, WW = wastewater, GW = groundwater, CSOs = combined sewer overflows.

Name	Attribute	Units	Description
Intergenerational equity			
rehab	% Realization of the rehabilitation demand	[% realization]	Each year, some parts of the water supply and WW system reach the end of their lifespan If these parts are not rehabilitated (i e repaired, renovated or replaced), negative effects concerning the performance and reliability of the WW system will likely occur Rehabilitation includes repairing damaged parts, renovation and thus an extension of the lifespan of assets (renewal), and their replacement
adapt	Flexibility of technical extension or deconstruction of infrastructure	[% flexibility]	Centralized systems (e g WW discharged in one single sewer network that reaches a central WWTP) are strongly connected and thus show strong path dependencies Decentralized systems (e g water for households from rain and treated on-site) are often more flexible In case of major changes, large adaptions of central systems are required, while decentralized systems may require addressing merely specific assets
Protection of water and other resources: Surface water			
chem	% Reference points in catchment that fulfill water quality target (nutrients, micropollutants, value > 0 6)	[% > 0 6]	Micropollutants and nutrients can have unwanted effects (e g eutrophication) Each indicator is classified into a quality class at each reference point (Bundi et al , 2000; FOEN, 2010; http://www modul-stufen-konzept ch) To aggregate all indicators, we use MCDA: the quality classes ("bad", "unsatisfactory", "moderate", "good", "very good") are translated to a neutral value (0 – 1) and aggregated to a value for that reference point (Langhans et al , 2013; Langhans and Reichert, 2011; Schuwirth et al , 2012b) For the catchment, we give the% reference points that fulfill the requirements (value >0 6)
hydr	% Reference points in catchment that fulfill VSA guidelines for stormwater handling	[% yes]	If too much rain water is discharged, a river may have hydraulic stress (turbulence, e g fish eggs in plants are washed out) The guidelines of the association of Swiss WW and water protection experts (VSA, 2002) are based on the ratio between the water amount coming from the river and from the discharged rain We give the% reference points in the catchment that fulfill the VSA guideline
Protection of water and other resources: GW			
gwhh	% Water abstraction / GW recharge	[%]	The GW resources in the case study are continuously replenished by percolation of rain water through the upper soil strata (GW recharge) But GW is also abstracted, e g as raw water for water supply If the abstraction exceeds the GW recharge, seasonal or permanent drawdown of the GW table is possible Sealing of the soil surface, especially in urban areas, further reduces recharge
exfiltrsew	Water quality class (of nutrients)	5 classes	Pollutants can reach GW from the WW system The probability of contamination increases when the sewers' condition decreases (e g many cracks) This attribute is based on expert estimates about the condition of the sewers, the% WW that infiltrates into the GW, the recharge rate of the GW and the nutrient indicators used for the "good chemical state of the watercourse"
exfiltrstruct	Water quality class (of biocides)	5 classes	Infiltration of rain from impervious areas (e g roofs, streets, parking lots) increases the risk of contaminating GW with pollutants For example, biocides are contained in building materials (to limit plant growth) The probability of contamination increases if more rain water is infiltrated via infiltration structures, if the pollutants are not retained as expected or if the infiltration structure is not built properly As indicators we use biocides
Protection of water and other resources: Efficient use of resources			
phosph	% Recovery of phosphate from WW	[% P recovery]	WW contains nutrients, e g nitrogen and phosphorus Many nutrients are mainly contained in human urine (not feces); e g about 80 % of nitrogen and 50 % of excreted phosphorus (e g Larsen et al , 2009; 2012) We use recovery of phosphate as indicator, because nitrogen is available in large amounts in nature Phosphate is vital as fertilizer for agriculture, but is limited to phosphate mines that will at some point be exhausted
econs	Net energy consumption for water / WW treatment and transport	DW: [kWh / m^3] WW: [kWh / p / yr]	In the technical DW and WW system, electrical energy is needed for e g treatment installations (aeration) and pumping The amount depends on e g the amount of water, treatment option and operating conditions The generation of electrical energy consumes natural and depletable resources (fossil fuels), which can emit greenhouse gases However, energy can also be recovered from the WW system
Good supply with water: DW: Good quality			
aes_dw	Days / year with esthetic impairment	[d / yr]	Occurring separately or together, smell, taste, discoloration and turbidity lead to esthetic impairment of DW The esthetic water quality depends on the raw water composition and

Name	Attribute	Units	Description
	e g taste, smell, etc		technical installations That is, the way and quality of raw water purification, dimensioning of pipes and reservoirs (stagnation), condition and maintenance of the distribution system (e g sediments, corrosion, microbial contamination)
faecal_dw	Days / year with hygienic concerns (hygiene indicators)	[d / yr]	According to the legal requirements, DW should be free of any hygienically unsafe organisms However, their occurrence is not impossible Causes for an occurrence include inadequate treatment, long stagnation of water in the supply mains and service connections, improper or inadequate cleaning of equipment or contamination from improperly connected pipes
cells_dw	Changes in total (cell count as indicator of bacterial re-growth)	[log]	Cell counts are an indicator of the amount of microorganisms in the water and thus help monitoring microbial regrowth in the supply network (changes in cell concentration) Most occurring (active) cells are harmless A distinction between active and inactivated cells, e g after disinfection, is not possible Although low cell counts are generally preferable to high cell counts, the absolute number of cells is not a direct indicator of water quality because each system has its own equilibrium concentration of cells
no3_dw	Inorganic substances (nitrate concentration)	[mg / L]	Nitrate is not toxic to humans in natural concentrations, but is regulated because of toxic (nitrite) or carcinogenic byproducts (nitrosamines) In Switzerland, the DW threshold value (40 mg / L) and the Water Pollution Control Ordinance ("Gewässerschutzverordnung") require 25 mg / L, mainly because of ecological concerns
pest_dw	Pesticides (sum of pesticide concentration)	[µg / L]	The sum of pesticides is an indicator of the overall burden of chemicals used in agriculture and in settlements To avoid adverse effects on human health and the environment, the Swiss DW directive states a maximum threshold concentration of 0 5 µg / L for the sum of pesticides and 0 1 µg / L for individual substances
bta_dw	Micropollutants (indicator: benzotriazole)	[ng / L]	Benzotriazole is an indicator of the overall burden of micropollutants in DW It is used in coolants, for corrosion protection and de-icing It is not removed by most standard treatment processes and one of the most ubiquitous environmental micropollutants in Switzerland Recommendations for discharge threshold concentrations from WWTP exist (120 µg / L for single; 30 µg / L for chronic discharges), but thresholds for toxicological concern in DW are under discussion
Good supply with water: DW: High quantity			
vol_dw	Days per year with water quantity limitations	[d / yr]	Of the water used in households, DW quality is strictly necessary for drinking, cooking, washing of food or dish-washing by hand, etc The proportion of DW in the kitchen of a Swiss household is estimated at 15 % A limitation occurs if the dimensioning of the system or the climatic conditions do not allow for the abstraction of more than the average yearly demand, that is, coverage of peak demands cannot be warranted
Good supply with water: DW: High reliability			
ci_dw	Criticality index (criticality of affected pipe x probability of outage / total criticality of all pipes)	-	Criticality is a dimensionless index between 0 and 1 which describes how many supply interruptions of what degree are to be expected due to technical system failure(s) To each water pipe an index is assigned that describes how critical this line is for the functioning of the system The larger the diameter, the greater the potential damage and the number of people who are affected by supply interruptions and thus, the higher the index for this pipe
Good supply with water: Household water			
	Same objectives and attributes as "DW"		Household water is that part of the water used in the household for anything except potable use This covers personal hygiene, toilet flushing, clothes washing, house-cleaning, etc
Good supply with water: Water for firefighting			
vol_ffw	Available water for firefighting in new housing areas	[L / min]	The current requirements of the Cantonal Building Insurance for firefighting protection assume that 1'500 to 3'600 L / min of water are needed during 30–100 minutes (firefighting reserve in central water reservoir) to extinguish a fire, depending on the type of building and its use Compliance with other rules on the maximum distance to the nearest fire hydrant and a minimum pressure of 3 5 bar in the distribution system are accounted for in the alternatives' design
ci_ffw	Same as for "Drinking and Household water"		
Safe WW disposal: Hygienic drainage and discharge			
illn	% Of total population getting infected once per year	[% / yr]	WW contains pathogens (bacteria, viruses), which may infect people and cause illness Gastrointestinal infections can lead to stomach disorders, diarrhea or vomiting Direct contact with WW is possible, if professionals or people that maintain private decentralized treatment plants do not follow precautionary measures It is also possible when the sewer system is overloaded and WW floods e g into cellars
cso	Number of combined sewer overflows (CSOs) per year per receiving water	[no / yr / receiving water]	Combined sewer systems carry WW and storm water in the same pipeline(s); it is treated in WWTP If the system is overloaded under heavy rain, the mixed WW is released directly to receiving waters without or only basic treatment (= CSO) If people go swimming or bathing in rivers or lakes after heavy rain with CSOs, there is a risk of getting infected with pathogens from WW, which may cause illness
Safe WW disposal: High reliability of drainage system			
failure	Weighted (by pipe diameter) no pipe collapses, blockages / year / 1'000 inhabitants	[no / yr / 1'000 people]	If sewers break down or are blocked, rain and WW are no longer effectively drained away from houses and streets Damaged sewers may collapse; if they are not well maintained, they may get blocked (e g plant roots, debris) Rain and WW that spill into streets or houses, or overflow parking lots may be a nuisance, can disturb traffic and business or can damage private or communal property
service	Weighted (by city center, inhabitants) number of incidents	[no / yr]	Even well-maintained urban drainage systems are not designed to provide undisturbed drainage during extreme rainfall Combined sewage and stormwater are then pressed to surfaces via manholes or enter cellars Floodings are a nuisance, can disturb traffic and business or damage

Name	Attribute	Units	Description
	of insufficient drainage capacity / year		property An incident of insufficient drainage capacity that affects a bigger area and historic town centers with mixed living and commercial zones is given a higher weight
High social acceptance			
auton	% of the water coming from the region Mönchaltorfer Aa	[%]	For some, access to their own water and water rights are important Water that is withdrawn by the water suppliers in the case study region or its immediate neighbors in the same geographical basin (Uster, Maur, Wetzikon) is classified as water from the region Water from lake Zürich is classified as outside the region
efqm	% Score of EFQM Excellence Model	[%]	The EFQM Excellence Model (formerly European Foundation for Quality Management), is the most popular quality tool in Europe, used by more than 30 000 organizations to improve performance It is used by firms for (self) assessment, an exercise in which an organization is graded against a detailed set of 9 criteria
voice	Degree (percent) of co-determination	[%]	Co-determination defines how and how much citizens can take part in the decision-making of water infrastructure planning It depends upon 3 factors: a) organizational structure of WW system and its legal form, 2) geographic extent of community and WW disposal utility and 3) financial strategy
time	Necessary time for operation and maintenance by end user	[h / person / yr]	In conventional central water supply and WWTP, professionals are responsible for the operation of the system and they carry out all the maintenance End users do not have to invest any time If treatment options are installed in households, at least one person in this house has to do some maintenance or contact an expert or hotline in case of malfunctions
area	Additional area demand on private property per end user	[m^2 / person]	Decentral water or WW treatment units are installed directly at the end users' location For this, they must provide space on their private property (e g in the cellar or garden) For example, low-tech decentralized options for WW are constructed wetlands, where the sewage water is lead into a planted field
collab	Number of infra-structure sectors that collaborate in planning and construction	–	Different infrastructures share the underground: transportation, gas supply, energy supply with district heating, telecommunication, DW supply and WW disposal The higher the number of suppliers that collaborate when they are planning measures to open the underground, the less unnecessary road and construction works are expected, with the respective consequences such as construction sites, noise and traffic congestion
Low costs			
costcap	Annual cost / person in% (DW) / CHF (WW) of mean taxable income	DW: [% / p / yr]; WW: [CHF / p / yr]	The total costs for water supply and WW infrastructures include running costs (operation, maintenance) and capital costs (investments, depreciation) for treatment plants and pipe systems; or for decentralized units The costs are discounted with help of a discount rate ("projected") to the end of the time horizon, until 2050
costchange	Mean annual linear increase of costs in% (DW) / in CHF (WW) / person / year until 2050	DW: [% / p / yr]; WW: [CHF / p / yr]	Generally, large increases of the costs from one year to the next are unfavorable, because either high amounts of capital are required or because large credits are needed to unfavorable interest rates We consider the average increase of the annual costs in% (DW system) or in CHF (WW) per person and year

2.3.2 Future scenarios

Three future scenarios were created in the workshop, which characterize a plausible world in the region "Mönchaltorfer Aa" near Zurich in the year 2050. The scenario "Boomtown Zurich Oberland" ("Boom") was based on massive population growth and high prosperity. "Doom" depicted a difficult situation of Switzerland and Europe in the global world, with a slight population decline and few resources for the water sector. "Quality of life" assumed qualitative growth and emphasized sustainable development (Tab 2; Lienert et al., 2014a). The scenario "Status Quo" was not developed in the workshop; it was essentially a long-term projection of the current situation (i.e. current population, finance etc.).

Table 2: Summary of four future scenarios for the year 2050. Details see Lienert et al. 2014. WWTP = wastewater treatment plant, DW = drinking water, WW = wastewater.

Scenario	General characteristics	Water sector
Status Quo (as 2010)	24'200 inhabitants in 4 rural communities[(a)] near Zürich Extensive agriculture Urban growth pressure Lake for leisure activities, nature protection zones Eutrophication problems	Fragmented water governance: 3 WWTP, several water suppliers High quality of DW Water usage ca. 215 liter/ person/ d (including small businesses; only household water: 135 liter/ person/ d) [(b)] Insufficient water quality in rivers receiving WW; contains micro- and other pollutants
(A) Boomtown Zürich Oberland	Highly prosperous region 200'000 inhabitants Dense urban development Lake Greifensee is nature protection zone New transportation axes (magnetic	High-tech water treatment, new technologies (on-site) Overall increased water demand, but lower per person usage [(c)] DW quality as today WW quality higher than today (remove micropollutants)

Scenario	General characteristics	Water sector
	levitation train)	
(B) Doom	Switzerland and Europe lose attractiveness, globally Strong financial pressure on water infrastructures Slight population decline Strong urban sprawl Decline of industries Communities have to collaborate	High DW demand [c] (162 liter/ person/ day household use; -25 % WW discharge) Very bad state of infrastructures Population uses own sources (bottled water, rain water) Increasing environmental effects due to low WW treatment Deficient urban drainage; climate change effects (floodings)
(C) Quality of life	Highly prosperous region Moderate population growth (<5 % / y, until 2050 ca. +20 % = 29'000) Only 5 % new building areas Good financial situation High environmental and health awareness	Higher DW quality Lower water demand per person [c] Public network, rain retention basins, advanced treatment ponds Very high quality standards for WW treatment Nutrient reuse from WW

[a] The communities are: Egg, Gossau, Grüningen, and Mönchaltorf.

[b] 215 liter water usage/ person/ day based on average water consumption for households *and small businesses* from 2008–2011 in case study communities. In the alternatives, we based our consumption estimations in *households* on statistical data from Switzerland and Austria (see attribute description in Lienert et al. 2014).

[c] Although some groups defined the exact water amount per person and day for their scenario, we did not use these, because water usage also depends on the alternative and because we based later calculations on different assumptions for the "Doom" scenario (see Methods).

These scenarios provided valuable input, but we had to further process them. For the "Boom" scenario with massive population growth (eight times current population until 2050), the workshop participants had spatial planning ideas (e.g. 25-storey skyscrapers; see Lienert et al., 2014a). We later carried out quantitative, simplified spatial planning with two other NRP 61 projects; iWaQa (2013) and AGWAM (2013) to better correspond to likely urban expansion in Switzerland[8]. We also modified the water demand (water usage/ person/ day)[9].

2.3.3 Step (3) Develop alternatives

Ten strategic decision alternatives were created in the stakeholder workshop (Tab. 3). These were combinations of various technical infrastructure options (e.g. central vs. decentralized treatment), maintenance and rehabilitation strategies (e.g. continuous replacement vs. no rehabilitation) and management aspects (e.g. public vs. privatized organizational forms). After the workshop, the project team put considerable work into specifying the alternatives and ensuring internal coherence. We had to create some new factors to distinguish between alternatives. These specified the detailed water and wastewater treatment technologies and several characteristics regarding organizational activities and quality enablers to assess the "% score of EFQM Excellence Model" attribute (Tab. 1). A narrative of each alternative based on the stakeholders' inputs and the factor specifications are given in Lienert et al. (2014a). We developed some additional variants, especially based on the Status Quo.

8 We based planning on Swiss standards, preserving agricultural land and forests. We used typical building features in dense areas of Swiss cities (Zurich, Geneva), with up to 10-storey houses, and allocated these to areas foreseen for urban development in the current spatial plans of the study region. We added additional building sites for the Boom scenario and increased the population to 200'000 without "building" skyscrapers.

9 The predictions for water demand are a function of scenario and alternative (e.g. water saving by using rain water or urine-separating toilets). Halving the water demand in the Doom scenario as defined in the workshop, for example, still translates into high water provision for the utilities, since there will likely be large water losses caused by low maintenance (leaky pipes).

Table 3: Summary of strategic decision alternatives (see Lienert et al., 2014a). WWTP = wastewater treatment plant.

No.	Alternative name	Description
A1a	Centralized, privatization, high environmental protection	Private firm provides centralized full-service for entire region; service as today but with micropollutant removal at WWTP (high environmental protection)
A1b	Centralized, IKA	As A1a, but provider is intercommunal agency (IKA)
A2	Centralized, IKA, rain stored	Intercommunal agency (IKA) provides centralized full-service, but rain is stored for firefighting
A3	Fully decentralized	Fully decentralized system in the responsibility of households with collection of rain water, bottled water from supermarket, and re-use of graywater
A4	Decaying infrastructure, decentralized in outskirts	Mixed responsibilities with minimal community service; decaying central infrastructures in core area, decentralized in outskirts; drinking water with POU [(a)] systems, or bottled water
A5	Decaying infrastructure everywhere	Community provides minimal service; cheap decentralized infrastructure in responsibility of households (as in outskirts of A4)
A6	Maximal collaboration, centralized	Maximal collaboration in a cooperative that provides centralized full-service; micropollutant removal at WWTP; strong focus on storm water retention
A7	Mixed responsibility, fully decentralized with onsite treatment	Cooperative and private responsibility; full decentralization; re-use of treated rainwater at POE [(b)]; on-site wastewater treatment; nutrient recovery for agriculture; storm water retention as A6.
A8a	Status Quo with storm water retention	Status Quo with storm water retention
A8b–A8f	Status Quo technical variants	Status Quo is modeled with different technical variants
A9	Centralized, privatization, minimal maintenance	Consumers choose private contractor that seek revenue-maximization; fully centralized system; minimal repairs only upon urgent need for action

[a] POU = Point of use treatment in the households to achieve drinking water quality; can be done e.g. on the tabletop or under the sink.

[b] POE = Point of entry (e.g. water treated close to where it enters household; at entry point from centralized water system or after water storage tank).

2.3.4 Feedback about the SDM-procedure

During each step of the SDM-procedure, we collected stakeholder feedback. We used this to critically analyze the main advantages and disadvantages of each step and give recommendations (Lienert et al.) (Tab. 4; details see Lienert et al. 2014).

Table 4: Summary of recommendations for the steps of the SDM-process, including advantages and disadvantages, based on own experience and stakeholder feedback. Details see Lienert et al. 2014. SH = stakeholders.

Step	Recommendation	Advantage	Disadvantage
1. Clarify decision context			
1.1	*Case study selection and delimitation of system boundaries*		
°	° Choose "real problem", i.e. SH need a solution	° High willingness of SH to participate	° Case study ≠ scientific project
°	° Clearly define interactions (type, number, length); look for support by important SH (as mediators) ° Strong commitment of researchers	° Increase willingness to participate ° Better knowledge about case study	° Lower flexibility to adapt to changes ° Mediators can be difficult to identify ° High time demand
1.2	*SH selection; clarify decision problem with SH*		
	° Stratified sampling (e.g. vertical axis: from local to national/ horizontal: engineering, administration, politics)	° Ensures broad coverage of SH	° Less obvious SH might be missed out
	° Face-to-face interviews (e.g. who plays role, is affected, interactions, interests, objectives) (specific: treat SH w. respect/ guideline/ creativity/ feedback/ simple language/ avoid scientific terms)	° Good representation of different perspectives ° In-depth knowledge about SH (e.g. interests, problems, interactions)	° Very time consuming (costly) ° Unrepresentative sample
	° SH selection with short questionnaire (Email, phone, internet survey): Who is important/	° Much faster procedure ° Broader (representative)	° Loss of in-depth knowledge

Step	Recommendation	Advantage	Disadvantage
	affected? Interests?	coverage	
	° SH selection with snowball sampling: Who else should we include? Who has very different view?	° Include specific knowledge of SH ° Include extreme perspectives	° Staying very close to initially chosen SH: all belong to same system
	° Ask for SH expectations (e.g. what is next step/ by whom? Expectations/ hopes/ fears/ recommendations?)	° Clarification: often SH expect practical outcomes (e.g. tool)	° Asking may lead to disappointment if expectations are not met
	° Clear communication/ information material about type of results and which expectations are/ are not met	° Avoids later disappointment if SH expect other outcomes	° Risk of disappointing SH at the start of the project
2. Define objectives and attributes			
2.1	° Set up objectives on desktop by research team, e.g. based on engineering requirements/ sustainability goals	° Objectives comply with methodol. requirements [(a)]/ state-of-the-art	° Loss of local SH knowledge ° Objectives may not meet SH needs
	° Face-to-face interviews; e.g. first open question ("what is fundamental?"); then consolidate w. existing objectives	° Avoids priming effects ° Focus on ideas/ objectives of SH	° Risk: too many/ diverging objectives ° Risk: ignore methodol. requirem. [(a)]
2.2	° Generate/ discuss/ consolidate objectives in workshop (e.g. brainstorming or present objectives from 2.1; discuss w. neighbor; discuss in plenum to seek consensus; use moderation methods to reduce number)	° Ideally, reflection of all opinions ° Better understand other SH opinions ° Bilateral gives voice to shy SH ° Ideally, focus on fundamental obj.	° Risk: objectives cannot be deleted ° Risk: no shared opinion ° Risk: ignore methodol. requirem. [(a)] ° Risk: lose control (moderation!)
2.3	*General recommendations for objectives and attributes [(b)]*		
	° <u>Understandability</u>: Use attributes common in field (weighting by all SH; elicit value function from experts)	° Based on scientific evidence ° Generalizable to other cases	° Technical/ natural-scientific attributes difficult for non-expert SH
	° <u>Missing or irrelevant objectives</u>: Generate objectives with intensive SH interaction (see 2.1, 2.2 above)	° If SH regards objective as irrelevant: give weight of zero	° Missing objectives cannot be added later; test sensitivity to this objective
	° <u>Attribute ranges</u>: define generalizable attributes; use relative numbers (absolute numbers for case study example; elicitation: avoid "range effect" bias)	° Allows for up- or down-scaling in other case studies	° Large ranges may be unrealistic ° Relative numbers may be less tangible/ more difficult to understand
	° <u>Preferential independence</u>: try to fulfill; check validity ("do preferences depend on levels of other attributes?").	° If this holds: simple additive aggregation model may be used	° If not given: more complex models needed (e.g. multiplicative)
	° <u>Minimum criteria</u>: some SH insist on minimal requirements, e.g. laws; discuss implication with SH	° Easy implementation in MCDA with minimum aggregation model	° Strong implications: exclusion of all alternatives not meeting minimum
3 Develop future scenarios			
	° Capture future uncertainty w.r.t. socio-economic development with snap-shot images. Must be very well prepared and moderated; convey that it is real science!	° Highly stimulating, very creative, fun ° Creates team-feeling; raises interest ° Invites thinking broadly about future	° Risk: not dealing with real problems ° Only limited participants possible ° Risk that things get out of control
4 Identify and create decision alternatives [(c)]			
	° SH workshop using creativity technique; e.g. create storylines of alternatives with scenarios as background	° Alternatives are relevant to SH ° Reduces anchoring on Status Quo	° Alternatives are not well worked-out: require further processing for MCDA
	° Combine creativity (above) with rigorous technique; e.g. strategy generation table (with/ without SH participation)	° All important elements are covered ° Internal consistency of alternatives	° Not very creative; tedious work ° Rather time consuming procedure

[a] Objectives should comprehensively cover decision, be fundamental, complete, concise, sensitive (distinguish between alternatives), non-ambiguous, understandable, simple, non-redundant, and be preferentially independent (to allow for additive value model). Based on feedback from later MCDA interviews for preference elicitation, not all requirements were met for all SH; see Lienert et al. 2014.

[b] Based on feedback mentioned in (a).

[c] Feedback concerning hypothetical alternatives and trade-off questions required in MCDA preference elicitation, see Lienert et al. 2014.

2.4 Discussion

In this paper, we developed a thorough, participatory procedure to support infrastructure planning processes in the water sector. Based on a real case study in Switzerland, we demonstrated how the initial steps of Structured Decision Making (SDM; Gregory et al., 2012a) can be carried out. Below, we compare our application with the general SDM-procedure. We discuss main advantages and disadvantages (Tab 4; details in Lienert et al., 2014a) before drawing conclusions.

2.4.1 Step (1) Clarify decision context

In environmental management, a single solution for a pressing problem is commonly sought after. However, also other decision types might be pursued such as "linked choices" or a ranking of risks (Gregory et al., 2012a). Sometimes, decisions will be repeated and an efficient, defensible decision system needs to be established. The here developed SDM-framework to support sustainable water infrastructure planning (SWIP, 2013) belongs to this type. The SWIP-approach must be transferable to other cases and accepted by those stakeholders. This is why we emphasized the first structuring steps of the SDM-process so much.

Selecting stakeholders is tricky even with the systematic approach that we followed. Gregory et al. (2012a) also acknowledge that it can be surprisingly difficult to determine decision makers. Even in one-off governmental (environmental) decisions, other stakeholders than the official representatives might have to be involved. It is usually unclear whether the participants of the SDM-process (Gregory et al., 2012a suggest five to twenty-five people) well represent society in general; often these are people with strong interests in the outcomes. We think that the general SDM-process of Gregory et al. (2012a) can profit from integrating tools for systematic stakeholder selection to ensure a good representation. We exemplified this in our SWIP-project with a detailed stakeholder and social network analysis (see Lienert et al., 2014a and references therein). This was based on 27 face-to-face interviews that lasted two to four hours each. Hence in our case, stakeholder characterization was connected with extensive effort, which absorbed over a year of the work of a PhD student (to set up interview guideline, find suitable participants, organize, carry out and transcribe interviews and analyze data). In most practice-oriented SDM-applications, we think that a short questionnaire survey among actors will suffice to ensure a fair representation of different perspectives (Tab. 4; details in Lienert et al., 2014a). However, this still entails more efforts than proposed by Gregory et al. (2012a).

Selecting the case study in research projects is typically driven by scientific considerations and less because a real-world problem needs solving (Renner et al., 2013). This also applied to our SWIP-example, but is problematic because it can hinder later collaboration with stakeholders (Tab. 4; Lienert et al., 2014a). In our case, we invested considerable time to convince stakeholders to collaborate. We had to rely on their goodwill to give us access to data or to participate in interviews after regular working hours. We therefore strongly recommend choosing a "real problem" as application, also in scientific projects. To increase the willingness to participate, we recommend defining the type, number and length of interactions. Moreover, often the expectations of stakeholders differ from what science can typically offer (e.g. Lang et al., 2012; Renner et al., 2013). For example, some of our interview partners expected a simple decision tool for infrastructure planning, which we cannot develop as part of this project (Tab. 4). To avoid disappointment, it is essential to ask about expectations and communicate the results that can or cannot be provided from the start.

2.4.2 Step (2) Define objectives and attributes

In the SWIP-example, generating the objectives hierarchy was extremely time-intensive. In the first desktop top-down procedure (Clemen and Reilly, 2001), the objectives were discussed in the monthly meetings of the scientific project team and intermittently processed during a year. The main advantage of this top-down approach is that we are sure that the objectives meet the engineering as well as SDM-requirements, e.g. that means objectives are avoided and that there is no double counting (see Methods; Tab.4; details in Lienert et al., 2014a).

We judge our approach to then consider individual stakeholder perspectives in face-to-face interviews as highly beneficial (see Methods, also concerning time demand). Hereby, personal viewpoints can be included on equal footing to consensus opinions. Priming effects can be avoided by using open questions. Interviews are not commonly used to generate objectives. In the literature, usually creative brainstorming-type stakeholder workshops are described, as recommended by Gregory et al. (2012a). The advantage of workshops is that fast agreement is possible (in our case five hours for the workshop plus a few days preparation). Structuring tools can support the workshop and visualization with e.g. means-ends networks is recommended (also see Clemen and Reilly, 2001 and Methods). However, workshops risk that fundamental aspects are missed because of too early consensus (the famous "groupthink" phenomenon; Janis, 1972; 1982 see e.g. review by Kerr and Tindale 2004)

Generating good attributes that are applicable to other cases again spanned several months of intermittent PhD student work. We had to use some "proxy attributes", which should be avoided (see Methods), but often we found "natural attributes", based on engineering considerations. Integrating environmental and societal objectives with traditional technical and economic indicators is a recent development in engineering (e.g. Ashley et al., 2008; Balkema et al., 2001; Lundie et al., 2004; Palme et al., 2005). We hope to contribute by presenting our attributes in detail (Tab. 1; details in Lienert et al. 2014). Additionally, we focused on the formal requirements for objectives according to decision theory, of which the engineering approaches such as Life Cycle Analysis (LCA) might have been less aware of. To make the proxy but also natural attributes more tangible, we combined them with narratives of the Status Quo and the best- and worst-possible case (Lienert et al., 2014a), as recommended by Gregory et al. (2012a). This was especially useful for the later elicitation of stakeholder preferences for the MCDA(Scholten et al., submitted; Zheng et al., 2013). We generalized the attributes if possible; e.g. the attribute of "good chemical state of watercourses" was set up together with the project iWaQa (2013) and covers the worst- and best-possible general case of water quality indicators (Schuwirth et al., 2012b). Similarly, "recovery of nutrients" from wastewater covers the whole range from 0 (as today) to 100 % (e.g. urine source separation and fecal collection; Larsen et al., 2009; 2012).

The aim of the SWIP-project to build a comprehensive, generalizable objectives hierarchy contrasts the condition of conciseness (e.g. Clemen and Reilly, 2001; Eisenführ et al., 2010; Gregory et al., 2012a; Keeney, 1992; Keeney and Raiffa, 1976). Gregory et al. (2012a) propose using only six to ten objectives, because people cannot keep track of more. If more seem necessary, objectives can be grouped into sub-objectives. We did this and built an objectives hierarchy with only six highest-level objectives (Fig. 1). Gregory et al. (2012a) further recommend context-specific rather than objectives for universal usage. The here proposed SWIP-objectives hierarchy (Tab. 1) is a compromise, since the intention is to support different, but specific decisions in water infrastructure planning, but not any general water management decisions. We encourage using our SWIP-objectives hierarchy, but advise others to carefully discuss excluding objectives that are irrelevant for their specific application. Moreover, the attribute ranges

(Lienert et al., 2014a) need to be adapted to other alternatives and system boundaries. If natural attributes are available in cases where we used proxies, these should be chosen instead.

2.4.3 Future scenarios

Scenario planning (e.g. Ringland, 2002; Schnaars, 1987; Schoemaker, 1995) is not a standard part of MCDA or SDM, but is recommended by Gregory et al. (2012a) to structure situations where it is difficult to assign probabilities. We found the combination of SDM with scenario planning highly fruitful. In our case study, the scenario workshop was that event that was the most fun for stakeholders. For example, one participant stated that: "It is great to step back from daily routine, question the current system and let your imagination run freely to think about the world in 2050." We made similar experiences in other scenario workshops (Lienert et al., 2006; Störmer et al., 2009; Truffer et al., 2010). We recommend stimulating workshops to get people "on board" and create a good project feeling. A risk of having fun is to imply that the project is not seriously dealing with the stakeholder's problems (Tab. 4; Lienert et al., 2014a). It is thus important to moderate workshops carefully. Workshops can only host few participants, but have the advantage of a fast generation of results. In our case, we invested a few days to prepare the five-hour workshop. However, the PhD students needed around four to five weeks to later specify especially the Boom scenario, which needed extensive "building" of new infrastructure. This was necessary for the following MCDA, but is not required if SDM is only used for structuring.

Scenario planning has recently entered the MCDA-literature. It is a tool to capture substantial external (scenaric) uncertainties for strategic decision making. In our project, we followed Goodwin and Wright (2001) who assume that the decision makers' preferences do not change under different futures. Hence, each stakeholder's preferences are elicited only once for the MCDA, instead of once for each scenario (e.g. Karvetski et al., 2009a; Lambert et al., 2012; Montibeller et al., 2006; Ram and Montibeller, 2012; Ram et al., 2011; Stewart et al., 2013). We argue that we have to make decisions today (about water infrastructures). These are grounded in the given state of the world, our subjective preferences and our ideas about the future. The preferences of each stakeholder include a subjective view about current conditions and a likely future. We emphasize that these individual future views are not captured by the scenarios: they do not contain a prediction[10] of what will happen, nor of the stakeholder's implicit beliefs[11]. If we use consistent preferences, we must ensure that the range of each attribute spans the entire possible (but uncertain) outcome under all scenarios (e.g. Stewart et al., 2013; ranges see Lienert at al. 2014). Thus MCDA-modeling efforts increase, because we estimate the performance of each alternative for each attribute under all four scenarios.

2.4.4 Step (3) Develop alternatives

There are many ways to creatively generate alternatives (see Methods). In our application, we combined a desktop approach with a stakeholder workshop. This ensured that stakeholders understand our methods, that alternatives are relevant to them and that they are later better accepted (Gregory et al., 2012a). It also avoids overlooking issues obvious to local practitioners. We also found the combination of the "strategy generation table" (Gregory et al., 2012b; Howard, 1988) with scenarios as background highly effective. In decision making, there is a very strong tendency to anchor on Status Quo alternatives (Nutt, 2004). We

10 "Second, scenario analysis usually tries to identify a set of possible futures, each of whose occurrence is plausible, but not assured. This combination of offering more than one forecast, and offering it in form of a narrative, is deemed by advocates to be a more reasonable approach than trying to predict (to four significant decimal places) what will happen in the future." (Schnaars 1987; p. 106).

11 The sum of the probabilities for the realization of the scenarios is not 1 but can be anywhere between 0 and 1.

thus expected stakeholders to create conventional infrastructure alternatives under "Status Quo", while the "Boom" scenario triggered high-tech on-site solutions and the "Doom" scenario cheap and simple alternatives (Lienert et al., 2014a). The strategy generation table then forced participants to rigorously cover main elements, thus contributing to the basic requirements of alternatives, i.e. internal consistency, completeness and comparability (e.g. Gregory et al., 2012a; Keeney and Raiffa, 1976). The strategy generation table well addresses this, but has the drawback of being tedious work. We thus recommend generating storylines in a creative stakeholder process, but that the project team provides the factor specifications (Tab. 4; Lienert et al., 2014a). The time demand was similar to the scenario workshop: preparation took a few days, but the PhD students invested three to five weeks thereafter to specify the detailed alternatives for the MCDA.

A further advantage of the strategy generation table is that it allows for fast screening of all imaginable strategies (Clemen and Reilly, 2001). It is e.g. possible to check whether each cell in the table is reflected in an alternative. This was mostly the case in our SWIP-example (Lienert et al., 2014a). It is recommended to iteratively improve or create new alternatives in SDM-processes (Gregory et al., 2012a). An important advantage of MAUT (contrary to outranking procedures) is that alternatives can easily be added later (e.g. Keeney and Raiffa, 1976).

Moreover, in our SWIP-application, feedback from later MCDA-interviews indicated that some stakeholders had difficulties to formulate preferences about unconventional alternatives (e.g. fully decentralized wastewater disposal; e.g. Guest et al., 2009; Larsen et al., 2009; 2012; Libralato et al., 2012b). Gregory et al. (2012a) discuss the opposite problem: that stakeholders suggest alternatives of which experienced environmental managers know that they do not work. They recommend including all proposed alternatives and using iteration in the SDM-workshops to check how well they perform. We followed this recommendation in our example and included all alternatives in the later MCDA. However, in the later interviews for preference elicitation, we explained the reasons for including unconventional alternatives; e.g. to provide general insights that surpass the daily problems in the case study (Scholten et al., submitted; Zheng et al., 2013).

2.4.5 Conclusions and Outlook

From the initial SDM-structuring steps (Gregory et al., 2012a) applied to the SWIP-case study, we can learn that e.g. the fundamental objectives "good water supply" and "safe water disposal" (Fig. 1) were undisputed among stakeholders. Feedback from interviews and workshops indicated that "intergenerational equity" and "high social acceptance" seem less important. Alternatives involving privatization or mergers (A1, A2; Tab. 3) perform especially well concerning "high quality of management and operations" and could help overcome deficiencies of the current fragmentation in the Swiss water sector (Dominguez et al., 2011; Lienert et al., 2013). Thus, the conventional solution (A8) might dominate alternatives A1 or A2. Negative aspects of decentralized solutions (time and area demand for end-users; e.g. A3, A4, A5, A7) are also characterized by objectives of "high social acceptance"; if these are not important, they could perform well. On the other hand, a positive aspect of decentralized alternatives is their flexibility: if "intergenerational equity" is unimportant, flexibility has little positive effect. Thus, we cannot conclusively dismiss e.g. decentralized alternatives at this stage, especially since their performance also depends on the scenario. A further analysis must follow the initial structuring phase.

In our case, we combine the SDM-framework with more-quantitative MAUT-analyses (e.g. Eisenführ et al., 2010; Keeney and Raiffa, 1976). Models are developed in our SWIP-project that predict the performance and decay of water supply and wastewater systems ((Egger et al., 2013; Scheidegger et al., 2011; Scheidegger et al., 2013; Scholten et al., 2013; 2014). We have elicited stakeholder preferences in second interviews (weights, single-attribute value functions, aggregation schemes, risk attitudes) and are currently carrying out MAUT-analyses (Scholten et al., submitted; Zheng et al., 2013). Hereby, we hope to identify one or several robust alternatives that perform well for most stakeholders under all future scenarios.

Adding a formal MCDA to the SDM-structuring process is one option. Obviously, our thorough SDM-procedure of "only" the initial steps is lengthy and expensive. In many cases, it may suffice to explore important issues by relying on elements of our work, which is why we present more details in Lienert et al. (2014a). For example, our objectives hierarchy (Fig. 1) and our strategic decision alternatives (Tab. 3) could be adapted in further analyses. New alternatives can easily be created with the strategy generation table (Lienert et al., 2014a). An engineering firm might estimate the performance of the decision alternatives, based on our attributes (Tab. 1). Once this information is available, one might put more resources into the most-promising alternatives.

We hope to contribute to MCDA, but also to real decision making with this work. Hopefully, we can provide some guidance to engineers or community planners that are confronted with the "daunting challenge" (Milly et al., 2008) – in face of an increasingly uncertain future – of finding sustainable solutions for safe water supply and wastewater disposal, which is of vital importance to us all.

2.5 Acknowledgements

We thank the Swiss National Science Foundation for funding; within the National Research Program "Sustainable Water Management" (NRP 61, www.nfp61.ch), project number: 406140_125901 / 1. We thank Cristina Fritzsche, Peter Reichert, Florian Schnetzer, Jun Zheng and Anja Zorn for good collaboration and for data collection and the interview partners and workshop participants for their patience, support and valuable input. We thank Richard Michell for improving our English. We also thank two anonymous reviewers and the guest editors for their critical reviews which helped to greatly improve the manuscript.

3 Combining expert knowledge and local data for improved service life modeling of water supply networks

Lisa Scholten[a,b], Andreas Scheidegger[a], Peter Reichert[a,b], Max Maurer[a]

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

[b] ETH Zürich, Department of Environmental Sciences, Institute of Biogeochemistry and Pollutant Dynamics (IBP), CH-8092 Zürich, Switzerland

Environmental Modelling and Software (2013), 42, 1-16. http://dx.doi.org/10.1016/j.envsoft.2012.11.013

Abstract

The presented approach aims to overcome the scarce data problem in service life modeling of water networks by combining subjective expert knowledge and local replacement data. A procedure to elicit imprecise quantile estimates of survival functions from experts, considering common cognitive biases, was developed and applied. The individual expert priors of the parameters of the service life distribution are obtained by regression over the stated distribution quantiles and aggregated into a single prior distribution. Furthermore, a likelihood function for the commonly encountered censored and truncated pipe replacement data is formulated. The suitability of the suggested Bayesian approach based on elicitation data from eight experts and real network data is demonstrated. Robust parameter estimates could be derived in data situations where frequentist maximum likelihood estimation is unsatisfactory, and to show how the consideration of imprecision and in-between-variance of experts improves posterior inference.

Keywords

Scarce data, expert knowledge elicitation, expert aggregation, Bayesian inference, water supply network, service life modeling

3.1 Introduction

3.1.1 Challenge

The coming of age of the water infrastructure poses an increasing challenge for utility managers. One of the key issues is to assess the long-term development of network rehabilitation demand. The motivation is to ensure that sufficient funding is raised and appropriately allocated to achieve the foreseen level of service. As a result, the last decade of water infrastructure management has seen increased development, testing, and application of mathematical models in rehabilitation planning and network failure estimation (Alvisi and Franchini, 2010; Dridi et al., 2009; Eisenbeis et al., 1999; Fuchs-Hanusch et al., 2008; Kleiner and Rajani, 2001; Pelletier et al., 2003; Rajani and Kleiner, 2001).

3.1.2 Network rehabilitation and survival modeling with scarce data

Within these models, the expected service life of water supply pipes (also referred to as "pipe lifetime" or "pipe survival"), is inferred from historic failure or replacement data. A shortcoming of these models is that they are only applicable for rather well-kept and extensive data sets, which are not ubiquitous in many utilities, as for example in Switzerland (>50 % of the population served by utilities with < 10.000 customers, (SVGW, 2009)) where even the best documentation does not help to overcome the prevalence of short network length and thus small sample size.

Different strategies have been proposed to handle scarce data situations (i.e. situations in which model parameters cannot be identified or are too uncertain to be of use for practical rehabilitation planning):

Purely data-based methods. Renaud, De Massiac et al. (2009) tried to overcome the scarce data difficulty by amalgamating the data from a number of French water utilities to calibrate the model, but found that this did not result in models that were more effective.

Purely expertise-based methods. The survival model for rehabilitation prediction and its parameters or quantiles are directly elicited from experts, for example based on cohort survival (Herz, 1995, 1998). Even though the value of subjective expert judgment is largely unquestioned in practice, only few have proposed its use in water infrastructure engineering (Dridi et al., 2009; Korving and van Noortwijk, 2008).

Bayesian combination of subjective expert knowledge with data, e.g. (Dridi et al., 2009). Especially for small data sets, Bayesian inference might be advantageous over frequentist (purely data driven) inference. While the likelihood function is often discussed in detail, e.g. (Mailhot et al., 2000), the elicitation and influence of the prior probability distribution are rarely explored. This work does not only discuss the derived likelihood function for left-truncated and right-censored data, but also shows a meaningful procedure for quantitative expert knowledge elicitation and combination with locally available data. The performance of the Bayesian approach is compared to using a purely data-driven frequentist estimation, on the basis of different amounts of data.

3.1.3 Background on expert knowledge elicitation and aggregation

Regarding the elicitation of expert knowledge (referred to as "expert elicitation"), a wealth of publications and guidelines exist. Recent reviews on the role of expert knowledge in modeling and important elicitation aspects are the works of Krueger et al. (2012), Kynn (2008) and Low Choy et al. (2009). The former not only provides an overview about the formal use of expert opinion in modeling practice, but also a discussion about common critiques such as the definition of expertise, and the representativeness of experts. Kynn (2008) offers a critical review of the past decades of psychological research on expert elicitation, as well as relevant work from other fields. She concludes that over a decade of research into heuristics and biases has been almost completely ignored by the statistical literature on expert elicitation. (Interestingly, in the literature regarding elicitation of Herz's cohort survival model for pipe survival estimation, the wealth of publications on expert elicitation seems to have gone similarly unnoticed.) The latter work of Low Choy et al. (2009) comprises a review of applications of expert elicitation throughout the ecological literature. Apart from these, the reader is referred to the works of (Ayyub, 2001; Cooke, 1991; Cooke and Goossens, 2008; O'Hagan et al., 2006) for more in-depth information on the historical and theoretical background. Even though the effect of imprecision in the elicited data itself is sometimes discussed (O'Hagan, 2012; O'Hagan et al., 2006; Oakley and O'Hagan, 2007), it seems that the possibility of explicit elicitation of such imprecision has been overlooked in the past. The elicitation guideline developed hereafter includes the elicitation of imprecise estimates.

Consulting multiple experts can be interpreted as an artificial increase in sample size of the experiment (Clemen and Winkler, 1999) with the objective of getting an approximation to the intersubjective knowledge of the expert community rather than the subjective knowledge of a single expert (Gillies, 1991; Rinderknecht et al., 2012). A key decision is the way to aggregate this information into one single distribution, which is also reflected in numerous publications, e.g. (Ayyub, 2001; Clemen and Winkler, 1999; Cooke, 1991; Genest and Zidek, 1986; Jouini and Clemen, 1996; Kuhnert et al., 2010; O'Hagan, 2012; O'Hagan et al., 2006; O'Leary et al., 2009).

Following the categorization of Clemen and Winkler (1999), aggregation can be achieved by mathematical and behavioral combination. Unless a mutual consensus of the experts is envisaged, elicitation is performed on an individual basis and later mathematically aggregated. Mathematical combination approaches are often further subdivided into axiomatic (also named classical or pooling) approaches and Bayesian approaches (Clemen and Winkler, 1999; Cooke, 1991).

Many of the axiomatic approaches consist of linear pooling (e.g. simple weighted averaging) and differ only in the weighting of the elicited probabilities. Weights can be equal, or different for individual experts, e.g. assigned according to confidence levels or calibration. In comparative aggregation studies equal weighting performed reasonably well, though it was outperformed by more complex weighting rules in specific situations (Cooke, 1991; Cooke and Goossens, 2008). Clemen and Winkler (1999) conclude that simpler aggregation methods such as the simple equally-weighted arithmetic average perform just as well as more complex methods, a notion widely supported by others (Larrick and Soll, 2006; O'Hagan et al., 2006).

Therefore, two axiomatic aggregation methods were chosen for comparison. In approach A the differences between experts are considered to stem from the variability between different networks, whereas in approach B it is assumed that all experts refer to the same network (but the expert statements are uncertain and therefore different). The experts were assumed to be equally qualified, thus assigning equal weights.

3.1.4 Elicitation of the parameters from a multivariate survival function

When it comes to practical elicitation, it is often not possible to elicit the unknown probability distribution directly, but only the observable quantities (Lele and Allen, 2006; O'Hagan et al., 2006). This is because the experts can neither be expected to define a specific distributional form nor to estimate distributional parameters of possible functional models directly unless specifically trained. In the case of multivariate survival models, correlation between model parameters makes direct elicitation of parameters even more unreliable. To overcome this limitation, an approach to elicit the model parameters indirectly from experts' judgment on selected quantiles of the survival distribution was developed.

3.1.5 Goal and structure of the paper

The objective of this paper is to present an approach to overcome the scarce data problem in water pipe lifetime modeling by combining expert knowledge and local data.

The methodic contribution of this approach consists of

1. nomination of considered survival models (section 3.2.1),
2. a specifically developed elicitation guideline to obtain (imprecise) survival function quantiles from experts (3.2.2),
3. inference of a bivariate prior distribution of the survival function parameters from the stated quantiles,
4. mathematical aggregation of the experts statements into a single prior distribution for the survival model (3.2.3),
5. formulation of a novel likelihood function of the survival model for censored and truncated data (3.2.4),
6. frequentist parameter estimation under varying amounts of data (3.2.4), and
7. Bayesian updating of the expert prior using different amounts of data (3.2.4).

The utility data considered are summarized in section 3.2.5. In section 3.2.6, it is presented how possible sources of uncertainty were dealt with. Results (3.3) consequently cover the elicited expert knowledge, parametric model identification, most appropriate aggregation method, as well as the performance of maximum likelihood estimation and Bayesian inference in light of scarce data. Conclusions for its use in water main survival modeling are drawn in section 3.4.

As this work is interdisciplinary, the aim is to address readers from different professional backgrounds. Please bear with us for making aspects explicit which an expressed specialist might judge as banality. For linguistic convenience, experts will be male, the interviewer and the analyst female.

3.2 Methods

3.2.1 Choice of the survival model

The age t_i is the age of pipe i at the end of its service lifetime. As the service lifetime is different for every pipe it seems natural to model t_i with a random variable T. Because negative lifetimes are impossible, the distribution of T should support only positive values. T can be described by its probability density function p(t) or its survival function $S(t)$=P(T>t).

In practice $S(t)$ is described by a parametric function that ideally has a small number of parameters while being flexible enough to fit the data. Three parametric models that satisfy these requirements, and allow for a great variety of shapes, are the Weibull, lognormal and gamma distribution (Wayne, 2004).

These three models are used to infer the survival function from (a) the elicited expert knowledge expressed in the form of stated quantiles, (b) the available utility data with frequentist inference, and (c) a combination of both by Bayesian inference. The Weibull distribution is parameterized with $\boldsymbol{\theta} = (\alpha, \beta)^T$ such that $E(T) = \beta\Gamma(1 + 1/\alpha)$ and Var$(T) = \beta^2\Gamma(1 + 2/\alpha) - E(T)^2$, the lognormal with $\boldsymbol{\theta} = (\mu, \sigma)^T$, whereas E($T$) =$\mu$ and sd(T)=σ, and the gamma distribution with $\boldsymbol{\theta} = (k, s)^T$, so that E($T$) =$ks$ and sd(T)=ks^2.

For Bayesian inference a prior distribution for $\boldsymbol{\theta}$ is required. However, as experts cannot be expected to make reliable statements about the distribution of $\boldsymbol{\theta}$, an approach to elicit the distribution indirectly has been developed.

3.2.2 Expert Elicitation

A generic elicitation guideline, developed within the "Sheffield Elicitation Framework" (Oakley and O'Hagan, 2010), has been a major guidance for the design and adaptation of the elicitation procedure described below. Further details can be found in (Arreaza, 2011).

Minimizing cognitive biases

The elicitation guideline was developed keeping minimization of cognitive biases in mind. These are attributable to misunderstanding or discrepancies between the experts' responses and an accurate description of their knowledge (Spetzler and Stael von Holstein, 1975). The underlying research is comprehensively reviewed in (Ayyub, 2001; Cooke, 1991; Eisenführ et al., 2010; Kynn, 2008; Low Choy et al., 2009; O'Hagan et al., 2006). According to Kynn (2008), following the development of cognitive models to describe the encountered cognitive biases, three dimensions are categorized:

1. Internal consistency (and coherence) is mostly concerned with how well the experts' statements fulfill or contradict the laws of probability.
2. External consistency deals mostly with the ability of a person to control overconfidence in giving probability statements (calibration).
3. Self-consistency (reliability) deals with the variation in between statements when performing repetitive tests.

Often, not the biases or bias categories themselves, but rather more prominent heuristics leading to such biases are cited or even intermixed, e.g.(Kuhnert et al., 2010). Such heuristics are availability, adjustment and anchoring, and representativeness, originally reported and explained by Tversky and Kahneman (1974).

Several measures to avoid distortions to the elicitation data are described in (Cooke, 1991; Kynn, 2008; Low Choy et al., 2009), among others. From these measures the guidelines to minimize biases (bias category in parenthesis) were compiled:

a) making the desired mathematical implication of questions explicit (internal consistency, especially general additivity),

b) naming frequencies along with probabilities (internal consistency, especially conditional probability, and general additivity),

c) using tools (also visual) or checks during the elicitation procedure, and discussing possible incoherence with the expert to ensure the laws of probability are not violated (internal and external consistency),

d) ordering the questions in such a way that anchoring is avoided, e.g. non-sequentially as in bi-section method (external consistency),

e) calibration of the expert based on training questions which are related to the test questions (internal and external consistency),

f) using different trials, duplicated assessment and different encoding of questions (self-consistency, i.e. reliability of the results), and

g) assessing only non-tacit assumptions, i.e. not asking for extreme probabilities of distributions (internal consistency).

An accurate elicitation procedure including checks and repetition can also lead to the reduction of uncertainty in the stated quantities, whereas adequate preparation enhances consistency and reliability (Low Choy et al., 2009). Kynn (2008), Low Choy, O'Leary et al. (2009) and Oakley and O'Hagan(2010) furthermore emphasize the importance of motivating the experts to participate with diligence.

Quantile elicitation method

When it comes to practical elicitation, it is often not possible to elicit the unknown probability distribution directly, but only observable quantities (Lele and Allen, 2006; O'Hagan et al., 2006). Given that experts could not be expected to estimate the correlated parameters of a bivariate survival model directly, experts were asked to give estimates of the age until which a certain proportion of a pipe cohort is expected to last, i.e. the quantiles of the age distribution (e.g. *"How long does it take until 50 % of the members of this pipe group have been taken out of service?"*). This *quantile elicitation method* is equivalent to the analyst stating probabilities or relative frequencies of a cumulative distribution and the expert estimating the expected age at replacement, see (Rinderknecht et al., 2011) for more details. This is different from the typical application of the quantile elicitation method, because the expert does not state the quantiles of a distribution describing the expert's uncertainty about a quantity to be elicited, but the expert states point estimates of the quantiles of a pipe survival distribution. The elicitation of marginal distributions characterizing the expert's knowledge of all quantiles of the elicited age distributions would be much too demanding and would still leave the problem of missing information about the dependence structure of these marginal distributions.

The quantiles selected for the interview are, in this sequence, the 95 %, 5 %, 50 %, 75 % and 25 % quantiles, characterizing the tails, the position, and two easily interpretable quantiles in between, which allow for adjustment of the shape of the curve, respectively. Winkler (1967) also suggests this sequence in order to avoid anchoring and adjustment effects. The motivation for 95 % and 5 % quantiles as outer ranges is based on the reported limited ability of experts to correctly express the extreme tails of distributions, e.g. when asking for 99 % and 1 % quantiles (Alpert and Raiffa, 1982 in Oakley and O'Hagan, 2007). Experts might find it easier not to specify their opinion with absolute precision. Thus, the elicitation guideline was developed for both precise (point estimates) and imprecise values (stated intervals).

Selection of experts

To ensure enough diversity of opinion and expertise while at the same time avoiding redundancy of information (Ayyub, 2001), eight individual expert interviews, at a duration of approximately two hours each, were performed. The experts were selected following suggestions from the Swiss Gas and Water Association (SVGW / SSIGE). People from different parts of Switzerland with major experience in the fields of planning, construction, operation and maintenance of water supply networks were chosen. All of them carry a higher education degree. An overview of the experts and their specific qualification is given in Table 13, Appendix A.

Choice of pipe groups

It is effective to differentiate the pipe network by material and laying period for pipe survival analysis (Fuchs, 2001; Kleiner and Rajani, 1999; Roscher et al., 2008). Stratification based on other criteria such as diameter, pressure zone, soil conditions etc. is possible, but was not done because this might likely overtax the abilities of the experts and make the elicitation overly complicated. Though diameter can be a useful grouping criterion (Carrión et al., 2010) it was neglected in this study because the focus is on small networks in which diameter differences are small and diameters larger than 300 mm are generally rare. Additionally, the more stratification of data, the smaller the sample sizes for parameter inference. Out of thirteen possible pipe groups formed from material and laying period, five were chosen based on their frequency of occurrence in Swiss water supply networks, familiarity of experts with them, and the time these pipe groups had been in service. They are: grey cast iron (3rd generation only, GI3), ductile iron (1st generation only, DI1), asbestos cement (AC), steel (ST), and polyethylene (PE).

Pre-elicitation information

As preparation for the interviews, all experts were supplied with pre-elicitation information material at least three weeks prior to elicitation. Therein, the purpose of the study and background were stated, along with further information. The information covered the five pipe groups to be elicited, a rough scheme of the elicitation procedure, and suggestions on how to prepare for the interview. The experts were asked to provide feedback on the selected pipe groups. Furthermore, they were requested to thoroughly read through an elicitation example on the service life of an imaginary pump group with formulated questions and potential answers. The reason for pumps instead of pipes was to stay within the domain of the expert, while at the same time avoiding anchoring effects.

Elicitation procedure

First, an elicitation briefing is done. It includes *setting the scene* (purpose and procedure of the interview, expert's expertise, clarification of questions, selection of the four most familiar out of the five proposed pipe groups), *focusing* (characteristics of the pipe groups, motivation), and *training*. Goal of the training is to familiarize the expert with the question layout and to sensitize the experts to possible biases. The training example is the survival of women born in Switzerland in the year 1940 (for the reasoning behind this see Appendix A). Cross-checks with real data (Cordazzo, 2006) help to highlight specific features potentially leading to biases during pipe survival elicitation. Using a different domain for training avoids anchoring of the interviewees.

After this follows the main elicitation, for each pipe group separately. In the beginning, the experience of the expert with the specific pipe group is explored. Then the quantiles are elicited.

Quantities are roughly visualized using 100 paper clips (representing 100 % of the pipe group) and a paper sheet with a time bar. Experts are requested to disregard replacement because of initial laying failures (e.g. within the first year after laying), and replacements following managerial or other considerations not related to age or condition, such as coordinative ground works with other infrastructure providers. This helps to focus on technical lifespans and not on effects of different management decisions.

For a second round, the 75 %, 50 %, and 25 %- quantiles are re-elicited using bets, adjusting the stated ages until the expert is indifferent between the bets. This technique is used to confirm the statements by making the experts think differently about the quantities. After the bets, the elicited values are read out and confirmed with the expert. These checks and repetitions ensure that experts' statements are reliable, consistent and correctly documented. Lastly, the experts are asked for a qualitative description of the imaginary density curve (if possible), which in turn reveals whether it can be assumed unimodal.

Experts are requested to assess half of the pipe groups with imprecise estimates and the other half using precise estimates. At the end, experts are asked for feedback on the difficulty of the interview, and their preference regarding precise and imprecise estimates.

A more detailed description of the entire elicitation procedure is given in Appendix A and in (Arreaza, 2011).

3.2.3 Derivation of experts' priors of survival function parameters and aggregation

To combine the experts' statements and to construct an intersubjective prior distribution for the survival model parameters, axiomatic, equally-weighted pooling was chosen. Therefore, the elicited quantiles are considered as data to which the survival functions are fitted using nonlinear least squares regression. The resulting estimates and variance-covariance matrix of the survival function parameters were used to parameterize a bivariate lognormal distribution then called *expert prior*. The ages t_k that the expert assigns to the cumulative probabilities π_k are treated as dependent variables. The goodness of fit of the Weibull, lognormal, and gamma distributions can be directly compared based on the residual sum of squares (RSS) because they have the same number of parameters.

At least two aggregation options for combining the different experts into one intersubjective, general prior for Swiss water networks arise (terms are according to the classification of Gelman and Hill (2009) for hierarchical model regression):

Partial pooling: Fitting a distribution to each expert's estimates separately, and subsequently aggregating these distributions into one prior distribution.

Complete pooling of all experts' estimates before fitting a distribution to the stated quantiles.

With regard to pipe service life estimation, all experts are considered to be equally credible. They should thus receive equal weights. The possibility of correlation among experts caused by similar training or exchange of experience cannot be excluded (Jouini and Clemen, 1996) and is thus to be accommodated in the aggregated prior.

It is expected that for this example, partial pooling will be more appropriate than complete pooling, because the individual expert priors take better into account the expectedly different underlying environmental conditions in the experts domains throughout Switzerland.

Aggregation option A: partial pooling

The experts gained their knowledge on different water distribution networks whose deterioration is determined by diverse local conditions. Thus, dissimilar distributions can be expected to result from elicitation and fitting. To ensure these different conditions are accordingly reflected in the prior, not only the individual imprecision, but also the in-between-variance of the single experts' fitted distributions shall be considered. The procedure to obtain this prior is:

1. The inverse S^{-1} of the parametric survival function S is fitted to the ages quantified by the expert with a non-linear least squares regression for each expert separately:

$$t_{e,k} \sim S^{-1}\left(\pi_{e,k} | \exp(\theta_e^*)\right) + \varepsilon_{e,k};\ \varepsilon_{e,k} \sim N(0, \sigma_e) \tag{1}$$

Thereof for each expert e, $e = 1 \ldots E$, an approximate multivariate normal distribution $p_e(\boldsymbol{\theta}^* | \mu_e, \Sigma_e)$ for $\boldsymbol{\theta}_e^* = \ln(\boldsymbol{\theta}_e)$ is obtained with normal distributed error $\varepsilon_{e,k}$. Accordingly, the parameters of the survival distribution $\boldsymbol{\theta}_e$ (Weibull, lognormal or gamma) are lognormal distributed: $p_e(\boldsymbol{\theta} | \mu_e, \Sigma_e)$. If the expert stated intervals, both endpoints of the intervals are used for the regression.

2. The mixture of all E distributions $p_e(\boldsymbol{\theta} | \mu_e, \Sigma_e)$ can then be used as prior distribution:

$$p(\boldsymbol{\theta} | \mu_1, \ldots, \mu_E, \Sigma_1, \ldots, \Sigma_E) = \sum_{e=1}^{E} w_e p_e(\boldsymbol{\theta} | \mu_e, \Sigma_e) \tag{2}$$

where w_e is the weight of expert e and $\sum_{e=1}^{E} w_e = 1$.

This model is a mixture of the fitted individual prior distributions of the experts. It is sometimes referred to as the *density version of the linear opinion pool* (Genest and Zidek, 1986).

3. Because the mixture of the experts priors is likely to be multimodal, it is approximated (or rather smoothed) with a two-dimensional lognormal distribution $\tilde{p}(\Theta | \mu, \Sigma)$ that has the mean μ and covariance Σ, calculated as follows:

The (raw) moments of a mixture are the weighted average of the same moments of the component distributions (Frühwirth-Schnatter, 2006). Therefore the first moment (the expected value) of the mixture is

$$E(\boldsymbol{\theta}) = \sum_{e=1}^{E} w_e E_e(\boldsymbol{\theta}) = \sum_{e=1}^{E} w_e \mu_e = \mu \tag{3}$$

where $E_e(\boldsymbol{\theta})$ is the expected value of the distribution of expert e. The second moment is derived from the covariance of each component distribution:

$$E_e(\boldsymbol{\theta}\boldsymbol{\theta}^T) = \Sigma_e + \mu_e \mu_e^T \tag{4}$$

The second moment of the mixture is the weighted average of the second moments of the component distributions:

$$E(\boldsymbol{\theta}\boldsymbol{\theta}^T) = \sum_{e=1}^{E} w_e E_e(\boldsymbol{\theta}\boldsymbol{\theta}^T) \tag{5}$$

Thereof, the covariance of $\tilde{p}(\boldsymbol{\theta}|\mu, \Sigma)$ is calculated:

$$\Sigma = E(\boldsymbol{\theta}\boldsymbol{\theta}^T) - E(\boldsymbol{\theta})E(\boldsymbol{\theta})^T \tag{6}$$

This prior does not necessarily become narrower when more experts are considered. It might even become wider if new experts have gained their knowledge from different systems with other conditions.

Aggregation option B: complete pooling

Another approach is to pool the data beforehand to perform one single regression over all the data at once. This only makes sense if experts are considered as independent measurement devices and if they assess values based on experience from the same or very similar systems. In this case, the variance between experts is interpreted as measurement imprecision. Practically, aggregation consists of one single weighted non-linear least squares regression for all experts together:

$$t_k \sim S^{-1}\left(\pi_k | exp(\boldsymbol{\theta}^*)\right) + \varepsilon_k, \quad \varepsilon_k \sim N(0, \sigma) \tag{7}$$

Quantiles from experts who stated intervals estimates (two measurements per distribution quantile, endpoints of the intervals used) received half the weight compared to quantiles from experts who stated precise estimates (one measurement per quantile). From the weighted non-linear least squares regression a multivariate normal distribution $\mathrm{p}(\Theta^*|\mu_e, \Sigma_e)$ of the estimated parameters $\boldsymbol{\theta}^* = \ln(\boldsymbol{\theta})$ is obtained. Therefrom a log-normal distributed $\mathrm{p}(\boldsymbol{\theta}|\mu_e, \Sigma_e)$ is derived for the parameters $\boldsymbol{\theta}$.

Complete pooling has the effect that the more experts are asked, the smaller the uncertainty of the prior becomes. This is because the number of measurements (experts) increases while the number of parameters to be inferred remains the same.

3.2.4 Model parameter estimation

Likelihood function for left-truncated and right-censored data

The likelihood function of a model is required for frequentist and Bayesian parameter estimation. The likelihood function expresses the probability density to observe life-spans $t = \{t_1, \dots, t_N\}$ given the parameters $\boldsymbol{\theta}$.

As in most utilities, historic failure and replacement data have not been systematically documented in the studied utility until rather recently. This causes a left-truncation-right-censoring (LTRC) data scheme (Figure 2).

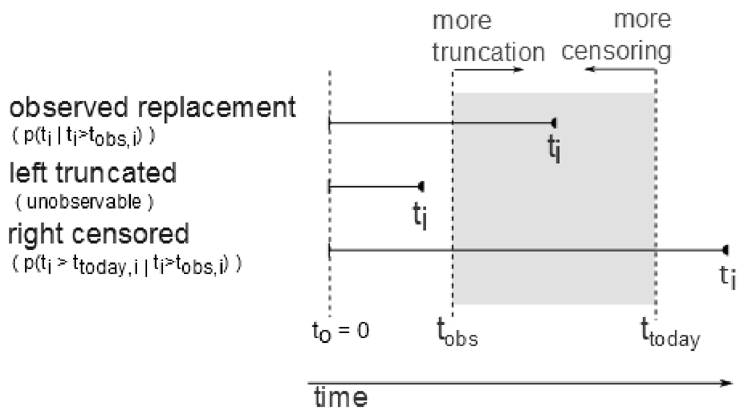


Figure 2: Left truncation and right censoring of a pipe group. The shaded area is the observation window between start t_{obs} and end t_{today} of observations.

Klein and Moeschberger (2003) describe right-censoring as an event which is only observed if it occurs before some pre-specified time, e.g. the end of a study. Consequently, pipes that are still in service at the end of the observation interval t_{today} are *right-censored observations*. *Left truncation* describes a situation where only subjects that have not yet experienced the event enter the study at a particular age and that are followed from this delayed entry time until the event occurs (or until the subject is censored). That means that data of pipes replaced before the start of observations are not available to the analyst, leading to only the more resistant of the pipes being observed.

The probability density to observe an uncensored age t_i of pipe i is written as:

$$p\left(t_i \middle| T > t_{obs,i}, \boldsymbol{\theta}\right) = \frac{p(t_i|\boldsymbol{\theta})}{S\left(t_{obs,i}|\boldsymbol{\theta}\right)} \tag{8}$$

where $t_{obs,i}$ is the age of the pipe i at the beginning of the observation period.

In situations where the end of lifetime could not be observed the likelihood for a single pipe becomes:

$$P\left(t_i > t_{today,i} \middle| t_i > t_{obs,i}, \boldsymbol{\theta}\right) = \frac{S\left(t_{today,i}|\boldsymbol{\theta}\right)}{S\left(t_{obs,i}|\boldsymbol{\theta}\right)} \tag{9}$$

where $t_{today,i}$ denotes the age of pipe i at the end of the observation period.

A censoring indicator δ_i allows for a short notation for the likelihood for all N pipes. δ_i equals zero if the datum is censored and one if uncensored. With this and the assumption that the pipes are independent the joint likelihood function for all pipes is:

$$\mathrm{p}(\mathbf{t}, \boldsymbol{\delta}|\, \boldsymbol{\theta}) = \prod_{i=1}^{N} \left(\frac{p(t_i|\boldsymbol{\theta})}{S(t_{obs}|\boldsymbol{\theta})}\right)^{\delta_i} \left(\frac{S\left(t_{today}|\boldsymbol{\theta}\right)}{S(t_{obs}|\boldsymbol{\theta})}\right)^{1-\delta_i} \tag{10}$$

This likelihood function is used for frequentist and Bayesian parameter inferences.

Frequentist parameter inference

Maximum likelihood estimation (MLE) is a common method to infer model parameters. The parameters that maximize the likelihood function for given data are used as best estimate. Large sample size properties of MLE allow the estimation of the variance-covariance matrix of the parameters from the inverse expected Fisher information matrix (Harrell, 2001), see Appendix B for more details.

Practically, this is done by a search through the parameter space by different optimization algorithms implemented in the R package *optimx* (Nash and Varadhan, 2011). The parametric models fitted are Weibull, lognormal, and gamma, as described in section 3.2.1. Multiple runs with different initial parameter values were performed to ensure stable estimates.

Bayesian inference

The aim of Bayesian inference is to update the prior probability distribution $p(\boldsymbol{\theta})$ with observed data $\{\mathbf{t}, \boldsymbol{\delta}\}$. The resulting posterior probability distribution is calculated with the Bayes theorem:

$$\mathrm{p}(\boldsymbol{\theta}|\mathbf{t},\boldsymbol{\delta}) = \frac{\mathrm{p}(\mathbf{t},\boldsymbol{\delta}|\,\boldsymbol{\theta})\mathrm{p}(\boldsymbol{\theta})}{\int \ \mathrm{p}(\mathbf{t},\boldsymbol{\delta}|\,\boldsymbol{\theta}')\mathrm{p}(\boldsymbol{\theta}')\,d\boldsymbol{\theta}'} \tag{11}$$

More in-depth information on Bayesian inference can be found in (Gelman et al., 2004). In this study, informative priors were derived based on expert elicitations and two different aggregation options (see sections 3.2.2 and 3.2.3). It can be shown that the choice of the prior distribution strongly influences the posterior result and is thus to be carefully chosen (Berger, 1990; Gelman et al., 2004).

The posterior distribution is derived by means of iterative Markov-Chain Monte-Carlo sampling (MCMC) with 6000 draws. The first 1000 draws are discarded as burn-in period and the acceptance rate is kept between 0.3 and 0.4 with the help of an adaptive sampler (Scheidegger, 2012; Vihola, 2011).

Graphical validation with a non-parametric survival estimate

Wayne (2004) suggests the use of a non-parametric survival estimate to visualize the fit of the parametric model. A Nelson-Aalen estimator adapted for LTRC data (also referred to as *extended Nelson estimator*) is used as described in Pan and Chappell (1998) and applied to pipe survival in Carrión et al. (2010). More details are given in the Appendix B.

The Nelson-Aalen estimator is capable of dealing with small sample sizes and can handle both censored and incomplete data as in our case of LTRC pipe survival (Klein and Moeschberger, 2003).

3.2.5 Utility data

The data used in this study consists of replacement records from a large Swiss water utility. Only pipe groups that were used in the prior elicitations were extracted from the provided pipe inventory. Reliable recording of pipe replacement started in this utility in 2000, so that only replacement entries between 2001-01-01 and 2010-12-31 were used for inference. The characteristics of the pipe groups are summarized in Table 5.

The effect of a decreasing amount of data (sample size) on parameter estimation is simulated by randomly reducing the available data to 500, 300, 150, and 50 pipes. These numbers correspond to the amount of data expected in mid-size or small water utilities for which Bayesian combination of expert opinion with local data is proposed. The ratio of replaced pipes to the overall number of pipes is kept constant. Shorter observation periods (more truncation, Figure 2) are also studied. They are accounted for by shifting the start of observation to 2003, 2005, 2007, 2008, and 2009 and removing replacement entries before this date, respectively.

Table 5: Summary characteristics of the examined pipe groups. Legend: GI3- 3rd generation grey cast iron (1930-1965), ST- steel, DI1- 1st generation ductile iron(1965-1980), PE- polyethylene, AC- asbestos cement incl. Eternit.

group	pipes (no.)	total length (km)	laying date (min-**med.**-max)	removal year (min-**med.**-max)	inner diameter # no. in mm	of which replaced
GI3	1295	104.5	1930-**1951**-1965	2001-**2005**-2010	[0- **100**[: 181 [100-300[: 854 [>300): 260	571 (**44.1 %**)
DI1	1009	87.45	1968-**1976**-1980	2001-**2007**-2010	[0- 100[:12 [100-300[: 865 [>300): 132	134 (**13.3 %**)
AC	153	22.45	1900-**1958**-1977	2001-**2006**-2010	[0- 100[: 31 [100-300[: 117 [>300): 5	38 (**24.8 %**)
ST	991	89.83	1875-**1965**-2009	2001-**2004**-2010	[0- 100[: 39 [100-300[: 594 [>300): 358	318 (**32.1 %**)
PE	195	18.02	1972-**2004**-2010	2006-**2008**-2009	[0- 100[: 25 [100-300[: 140 [>300): 30	6 (**3.1 %**)

3.3 Results and Discussion

3.3.1 Expert elicitation

The developed guideline was used for interviews with eight experts, numbered E1… E8. They provided estimates for grey and ductile cast iron (GG3 and DI1), whereas for steel (ST), polyethylene (PE), and asbestos cement (AC) only three, five, and six experts, respectively, provided their opinion. E4 estimated the service life of three pipe groups only, because in his network domain the two other materials were too rarely used. The obtained estimates are given in detail in Table 12 (see Appendix A) and summarized in Table 6. Two estimates were not considered in the later analysis: E1 for polyethylene, because the expert was highly uncomfortable about giving estimates for this material and could not imagine in which way it might deteriorate; E4 for steel pipes, because they were declaredly based on a past decision of this expert's utility to replace all steel pipes within only five years.

During the feedback at the end of the interview, five out of eight experts said they favored giving intervals instead of point estimates (Arreaza, 2011).

Table 6: Summary statistics for stated quantile values (ages) from experts for given cumulative probabilities. Mean = arithmetic mean of stated ages, sd= standard deviation. Pipe groups are explained in section 3.2.2.

group	GI3		DI1		AC		ST		PE	
probability	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
0.05	38.3	15.3	22.9	8.5	36.3	14.1	32.0	8.4	38.8	22.5
0.25	54.2	16.4	37.9	12.0	64.4	19.0	48.2	8.9	68.1	38.0
0.5	78.3	15.3	55.0	9.8	81.3	22.5	60.4	11.4	86.9	35.1
0.75	90.0	16.4	68.6	9.7	98.1	26.7	74.6	19.5	98.8	30.8
0.95	105.4	18.3	81.1	11.5	115.6	30.6	87.0	24.9	117.5	44.0

The summary statistics in Table 6 show that the quantile estimates between pipe groups, visible from the quantile means, are clearly different. With regard to the quantile standard deviations, not only a pronounced difference between materials is visible, but also an increasing uncertainty towards the upper quantiles. E4 and E5 gave distinctively lower estimates than other experts (Table 12 in Appendix A). Contrarily, estimates from E8 were consequently larger for all pipe groups. These visible differences between material groups and single expert values indicate the experts' awareness and ability to differentiate the aging behavior of the selected pipe groups. E4 and E5 named specific influences, such as strong deficits in laying or bedding quality, or difficult environmental conditions that could explain lower estimates (Table 11 in Appendix A). The longer lifetime suggested by E8 might also stem from anchoring to rather high values established by a former study this water supplier had commissioned. Other than this, the additional information given by the experts roughly allows us to explain differences between experts' statements and is thus considered as reflection of the encountered variability of conditions in the utility networks.

The usefulness of an expert is usually judged upon his contribution to an increase in knowledge. Measuring this usefulness based on the precision of statements or contribution to noise reduction, e.g. (Lele and Allen, 2006; Runge et al., 2011) is not appropriate in a case like ours. Rich knowledge is not necessarily equivalent to a high density of the mean and little spread of the fitted expert distribution. If the expert bases his knowledge on a variety of different water networks (or other objects of study), he might well accommodate this in more imprecise statements. Also, personal confidence and interrogation layout may play a role. An overconfident expert is likely to state shorter intervals than necessary to reflect his confidence levels (Speirs-Bridge et al., 2010). In this study, experts were encouraged to adequately consider their uncertainty in giving interval statements. The more useful expert is thus the expert stating wide enough intervals that contain his uncertainty about the quantity.

3.3.2 Parametric model identification

Non-linear least squares regressions were performed with the Weibull, lognormal, and gamma parametric models over the individual experts' statements. The goodness of fit measures were calculated (see residual sum of squares (RSS) in Table 15, Appendix B). The Weibull distribution provides the best fit for 23 out of the 28 single expert assessments. In one case, the RSS of the Weibull distribution is equal to the lognormal (E2 for polyethylene), and only inferior to the gamma or lognormal distribution in four single regressions (E8 for grey cast iron and polyethylene, as well as E5 in the case of ductile cast iron).

Similarly, maximum likelihood estimations (MLE) were done with the available pipe replacement data (see description in 3.2.4) for all three distribution models. Table 7 shows that the Weibull distribution for DI1, ST, and PE leads to smaller likelihoods than the lognormal or gamma distributions. In the case of GI3 and AC, the lognormal model fits the data slightly better.

Table 7: Obtained parameters and likelihood values of pipe group data from a large Swiss water supplier. Bold numbers indicate the model that maximizes the likelihood value log $p(t, \delta | \theta)$. Pipe groups are explained in 3.2.2 ("choice of pipe groups").

	Weibull distribution			lognormal distribution			gamma distribution		
	α	β	ln $p(\mathbf{t}, \boldsymbol{\delta} \| \boldsymbol{\theta})$	μ	σ	ln $p(\mathbf{t}, \boldsymbol{\delta} \| \boldsymbol{\theta})$	k	s	ln $p(\mathbf{t}, \boldsymbol{\delta} \| \boldsymbol{\theta})$
GI3	2.07	45.54	-2206.70	3.88	0.33	**-2204.74**	6.96	6.98	-2514.00
DI1	5.48	47.45	**-686.89**	3.86	0.31	-687.48	13.23	3.78	-687.13
AC	2.46	60.90	-172.59	4.07	0.29	**-170.50**	12.34	6.26	-171.27
ST	2.46	48.77	**-1259.96**	3.73	0.40	-1275.74	6.38	7.08	-1263.49
PE	1.81	65.72	**-36.29**	4.65	1.45	-36.82	1.85	58.73	-36.45

Though the Weibull likelihoods for GI3 and AC are only slightly larger than the lognormal, important deviations between the two models exist. This can be visualized by graphically comparing the nonparametric extended Nelson-Aalen estimation (see section 3.2.4, "Graphical validation with a nonparametric survival estimate") with the parametric models (Figure 3).

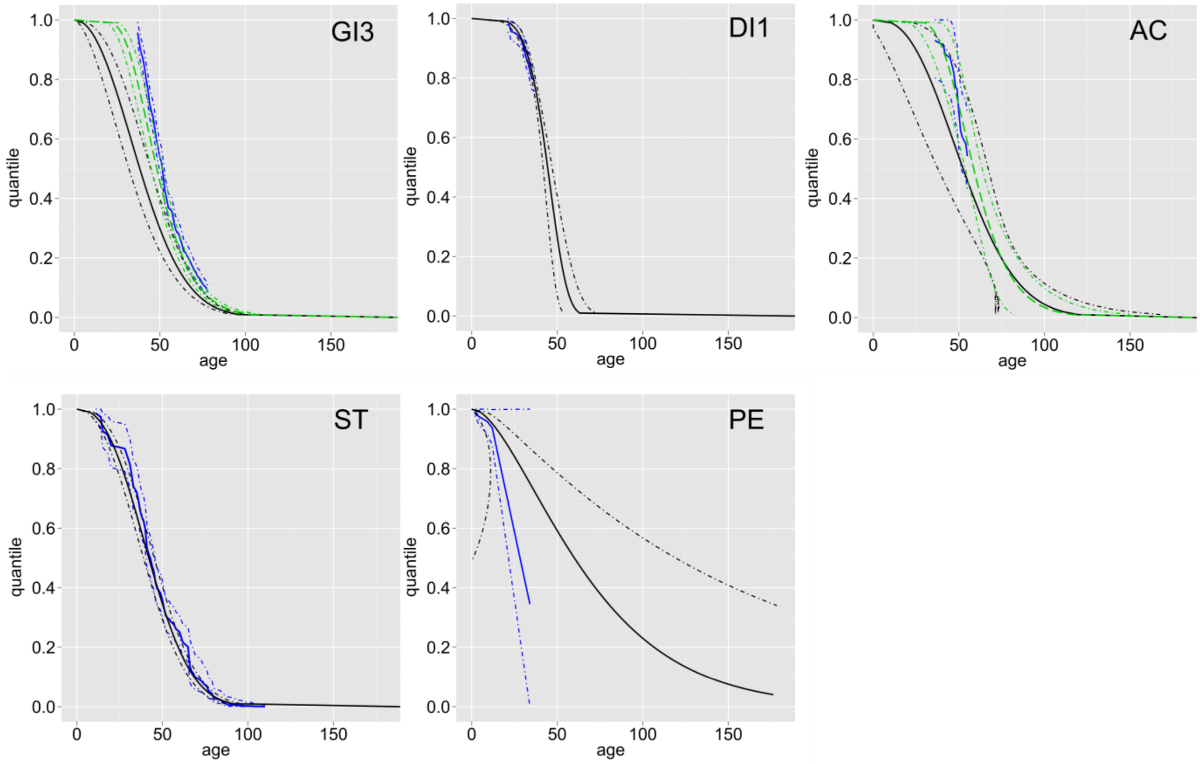


Figure 3: Comparison of a Weibull fit (black solid lines) and a nonparametric fit (blue step lines) for five pipe groups. The lognormal fit (green dashed lines) was added to the plot where it provided the best fit. The dash-dotted lines correspond to the 95 % confidence intervals of the mean (solid line).

The advantage of a nonparametric estimation is that it reflects the lifetime distribution more realistically within the observed time interval and can be used as a control for the possible parametric models. If both parametric and nonparametric models are overlaid, the parameter estimates satisfactorily approximate the „true" parameter values. Nevertheless, the nonparametric estimation model cannot replace parametric models because the outcome is a discrete estimate of the survival function (Coolen, 1996). It cannot be extended to unobserved intervals in time, something especially important for materials with a rather short history (for example for newer materials such as ductile cast iron or polyethylene). This is inconvenient for forecasting and especially for error propagation. Theoretical parametric models can be found that allow for extrapolation outside the data range and error propagation in combination with coupled predictive models. The structural deficiencies of the model are reflected in larger parametric uncertainty.

The Weibull model is best at approximating the nonparametric survival curves of DI1 and ST, but inferior to a lognormal distribution for GI3 and AC. Therein, the tail behavior of the lognormal distribution allows for an overall steeper survival curve whereas the Weibull model leads to underestimation of about 40 % to 50 % of the studied GI3 and AC cohort survival. The graph for polyethylene shows that the available data are clearly not sufficient to infer a trustworthy predictive distribution by means of MLE. This is also reflected in the uncertainty of the estimated parameters as specified in Table 8 (see MLE for all data), given a Weibull survival model for all materials. Despite the better fit of a lognormal model to GI3 and AC data, in the following sections regression and inference of parameters from the utility data is done for the Weibull model, unless otherwise stated.

3.3.3 Expert prior aggregation

Table 8 shows the mean parameter estimates of the Weibull shape ($\hat{\alpha}$) and scale ($\hat{\beta}$) parameters and corresponding uncertainty measures using a lognormal error distribution, $sd(\hat{\alpha})$ and $sd(\hat{\beta})$ respectively, for individual and pooled experts (see section 3.2.3). The parameters reflect the differences between pipe materials and experts as described for the elicitation results in 3.3.1. The parameter estimation of the complete and partial pooling prior results in similar means. Judging from a comparison of the aggregated scale parameters, the 63.2% quantile, AC pipes are believed to have the longest service life, followed by GI3, then PE, ST, and DI1 pipes. These observations are in line with survival estimates from the literature for German and Austrian water utilities (Fuchs, 2001; Roscher et al., 2005; Trujillo Alvarez, 1995), where AC and GI3 are usually judged most durable and DI1 least durable of the five considered pipe groups (Table 14, Appendix A). The impact of smoothing the experts' mixture on the prior used for inference is visible in Figure 4. The simple mixture according to equation (3) has a multi-modal density whereas the smoothed mixture, equation (4), is unimodal.

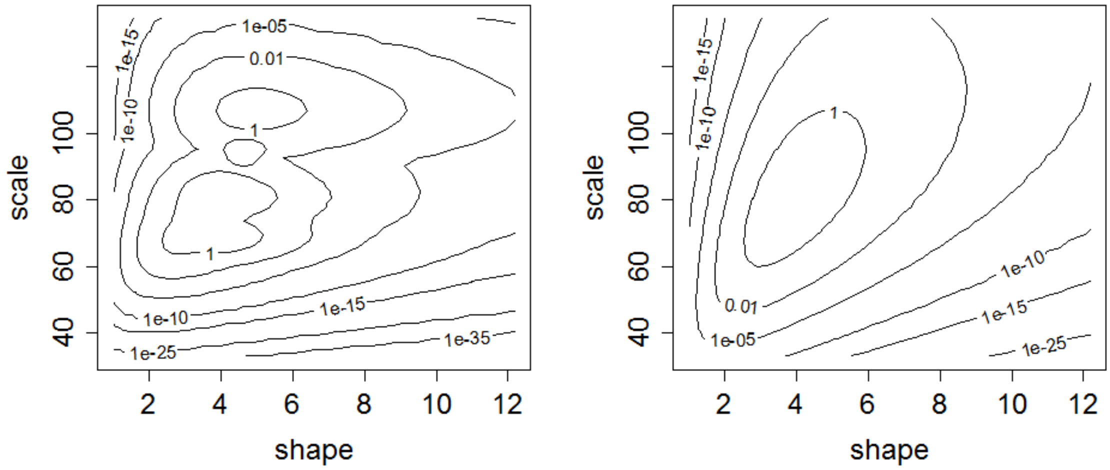


Figure 4: Bivariate probability density distribution of the aggregated prior (partial pooling) before smoothing (left, multimodal) and after smoothing (right, unimodal) for GI3.

Figure 5 shows the expert statements and mean survival function including 95 % confidence intervals for partial pooling as compared to the confidence intervals for complete pooling (exemplary for GI3). Unsurprisingly, the partial pooling yields larger standard deviations. This is because partial pooling incorporates the in-between variances, thus allowing for a better representation of the underlying differences in the experts' domains. The variance is not simply attributable to the experts' measurement error. As discussed in section 3.3.1 and 3.3.2, not only did the experts state diverse reasons for different aging behaviors in their utilities, these differences are also reflected in their statements. It is important to make clear that the experts do not have to agree in this context. If the expert judgments represent different distributions, of which each describes the underlying pipe survival in the expert's water utility, both in-between-differences and individual uncertainty are important sources of information. As opposed to this, complete pooling does not accommodate the in-between variance. It considers all judgments as measurements resulting from assessment of the same underlying distribution and its parametric uncertainty reproduces the error attributable to lack of fit from the model.

Although presumably obvious in a Bayesian learning framework, the aggregation approach chosen is not Bayesian. This is mainly owing to the need to formulate an unbiased prior to be combined with the individual priors, and that furthermore accurately considers the dependence structure between experts (Clemen and Winkler, 1999; O'Hagan et al., 2006). The definition of such a supra-prior is not only demanding, but also the analyst's independence from the experience gained during elicitation and data analysis cannot be expected.

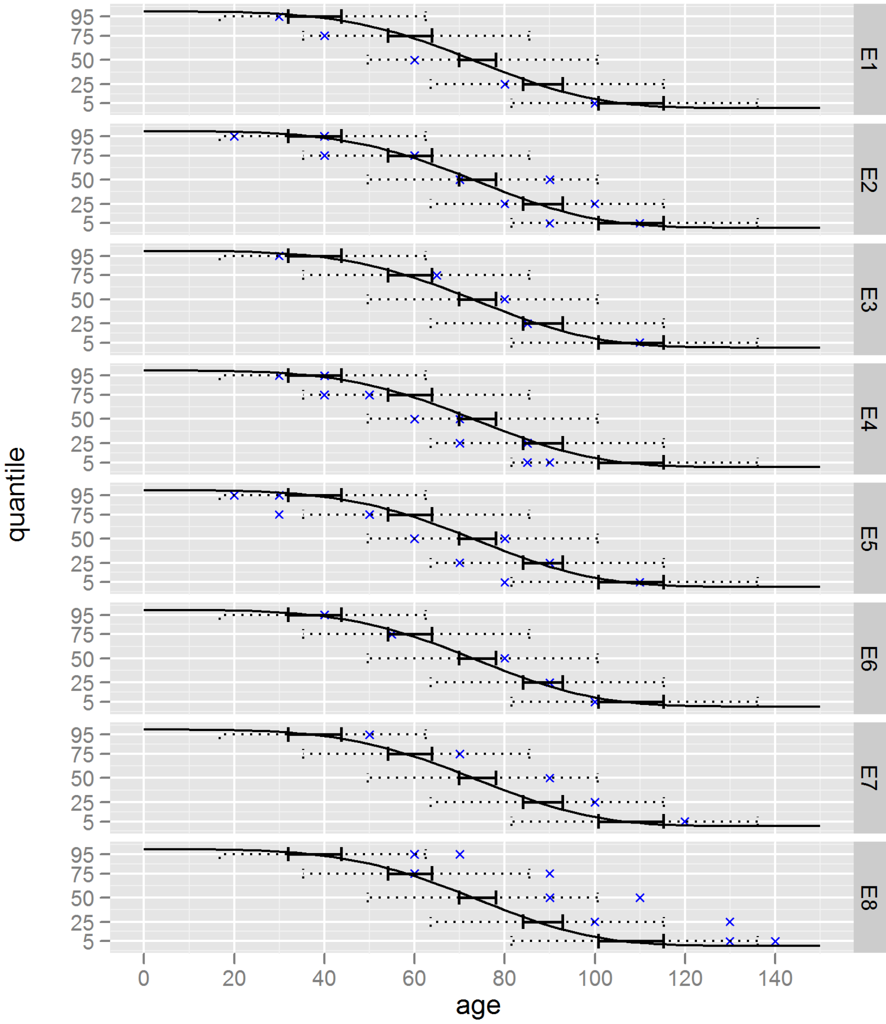


Figure 5: Comparison of GI3 priors and estimates from experts. Blue crosses represent quantile values as stated by the expert indicated on the right edge (E1…E8). Solid error bars give the 95 % confidence intervals for complete pooling, dotted error bars for partial pooling. The survival curve is calculated from the mean parameters ($\hat{\alpha}$= 3.97; $\hat{\beta}$= 81.22) of the partial pooling prior, see Table 8.

Table 8: Results from 1) non-linear least squares regression over experts E1…E8 (“Single experts”), 2) parameters obtained for the two aggregation methods complete pooling and partial pooling (“Aggregation”), and 3) maximum likelihood inference for all data, shortened observation windows, and artificial data reductions (“MLE”). The survival model is a Weibull distribution with parameters S($\boldsymbol{\theta}$)= $(\alpha,\beta)^T$.

		Grey cast iron (1930-64)				Ductile iron (1965-80)				Asbestos cement				Steel				Polyethylene			
		$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)
Single experts	**E 1**	2.96	69.16	0.28	2.18	2.04	57.15	0.19	2.40	3.49	99.24	0.51	4.27	-	-	-	-	-	-	-	-
	E 2	3.68	78.20	0.73	4.45	2.53	51.51	0.41	3.21	4.52	99.48	0.84	4.56	-	-	-	-	4.00	85.21	0.23	1.30
	E 3	3.77	82.40	0.53	3.26	4.05	66.02	0.53	2.30	3.87	110.9	0.42	3.27	-	-	-	-	-	-	-	-
	E 4	4.13	68.52	0.54	2.33	3.00	47.91	0.51	2.72	-	-	-	-	-	-	-	-	4.54	54.91	0.24	0.70
	E 5	3.18	69.62	0.64	4.50	2.97	58.00	0.40	2.60	2.93	55.99	0.14	0.93	-	-	-	-	2.38	44.62	0.22	1.67
	E 6	4.38	80.52	0.56	2.59	4.08	58.73	0.54	2.03	-	-	-	-	-	-	-	-	4.71	97.63	1.02	5.06
	E 7	4.61	94.43	0.21	1.03	3.44	60.42	0.41	2.17	3.95	63.90	0.64	2.80	3.67	73.25	0.50	2.87	-	-	-	-
	E 8	5.05	107.2	0.93	4.46	3.88	73.17	0.51	2.64	-	-	-	-	3.88	73.17	0.51	2.64	4.85	110.5	0.42	2.24
Aggre-gation	**complete pool**	3.92	81.31	0.36	2.05	3.16	59.34	0.22	1.30	3.79	86.01	0.60	3.83	3.75	73.19	0.34	1.84	4.22	78.78	1.02	4.90
	partial pool	3.97	81.22	0.91	12.70	3.25	59.11	0.88	7.83	3.75	86.05	0.74	24.14	3.77	73.21	0.52	2.76	4.11	74.40	1.21	26.73
MLE	**All data**	2.07	45.54	0.20	3.14	5.48	47.45	0.74	1.82	2.46	60.90	0.85	6.21	2.46	48.77	0.13	1.58	1.81	57.38	0.58	33.60
	≥ 2003	1.79	38.88	0.24	5.19	5.85	46.60	0.99	1.87	2.17	59.79	1.41	15.71	2.92	53.48	0.18	1.73	-	-	-	-
	≥ 2005	2.37	49.81	0.32	4.23	5.20	47.85	1.27	2.52	3.21	2866	0.52	2272279	2.97	55.45	0.24	2.13	-	-	-	-
	≥ 2007	2.49	49.81	0.42	5.27	3.98	51.10	1.96	4.34	-	-	-	-	2.79	51.95	0.30	2.84	1.21	145.06	0.64	359.62
	≥ 2008	2.37	52.85	0.59	8.03	4.68	50.45	2.64	4.89	-	-	-	-	2.96	53.52	0.39	3.40	-	-	-	-
	≥ 2009	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	500 pipes	1.95	43.81	0.33	5.88	3.66	53.94	1.08	5.24	-	-	-	-	2.42	47.68	0.18	2.21	-	-	-	-
	300 pipes	2.22	48.04	0.41	5.97	5.51	47.67	1.40	3.40	-	-	-	-	2.61	52.01	0.23	2.83	-	-	-	-
	150 pipes	2.91	55.32	0.65	5.53	3.24	60.22	2.44	13.82	1.81	56.25	0.99	16.51	2.37	46.70	0.30	4.07	2.39	39.29	0.70	13.71
	50 pipes	2.39	49.72	1.30	18.52	5.64	51.45	5.13	12.36	6.25	63.05	3.01	3.61	3.42	54.59	0.61	5.04	-	-	-	-

3.3.4 Maximum likelihood estimation from data

A maximum likelihood estimation (MLE) of the parameters of the survival function was done with the available pipe replacement data. The underlying model used is Weibull, the same model as used for describing the prior distributions of the experts. However, it must be noted that MLE of the parameters is independent of these priors.

For GI3, DI1, and ST, reasonably certain parameter estimates were derived (Table 8). Regarding AC and PE, a set of parameters was obtained as well, but with larger uncertainty due to the smaller number of data. The parameters from MLE show substantial differences as compared to the aggregated experts. Taking the Weibull scale parameter ($\hat{\beta}$) as representative for the age reached by 63.2 % of pipes in this group, basically the same ranking as given by the experts is observed: AC reaches higher ages than (in this order) PE, ST, DI1, and GI3. The exception is GI3, which is second most durable according to the aggregated experts, but least durable if only inferring from the data. Compared to the characteristic lifetime inferred by MLE from data, the aggregated experts estimates are approximately 11 (DI1) to 35 (GI3) years longer.

The results from randomly reducing the data to 500, 300, 150, and 50 pipes, while keeping the ratio of replaced to in-service pipes constant, demonstrate that the fewer data are available, the more uncertain the parameter estimates get. Analogously, increasing truncation when reducing the observation period to seven, five, three, two, and one year(s) leads to increasingly uncertain parameter estimates. This truncation is mirrored in the diminishing ratio of documented pipe replacements to pipes in service, given in Table 9.

Table 9: Effect of truncation on the ratio of replaced pipes to pipes in service. Ratios for which no MLE parameter estimates were obtained are highlighted.

	GI3	DI1	AC	ST	PE
2000	0.44	0.13	0.25	0.32	0.03
2003	0.36	0.12	0.18	0.23	**0.03**
2005	0.26	0.10	0.15	0.16	**0.03**
2007	0.18	0.06	**0.08**	0.11	0.05
2008	0.10	0.04	**0.07**	0.07	**0.00**
2009	**0.04**	**0.02**	**0.04**	**0.03**	**0.00**

For less than two years of observations, no parameters could be identified. As visible in Table 9, this can be attributed to ratios of less than ca. 10 % (GG3), 4 % (DI1), and 7 % (ST). Records of less than five years for AC (ratio of approx. 15 %), did not allow for any reliable parameter estimates, possibly explicable by the small sample size (153, Table 9). The scale parameters for PE and AC ($\hat{\beta}$) can only be estimated with very large uncertainty sd($\hat{\beta}$), an effect possibly caused by the small sample size and increasing truncation leading to a flat likelihood surface, such that the parameters are difficult to estimate.

Consequently, if a utility has only recently started reliably recording pipe replacement (high truncation), or if the number of pipes in the network is small, no reliable parameters can be found with MLE.

Furthermore, the considerable differences between the distribution parameters inferred from expert statements and utility data cannot merely be attributed to more extreme local conditions, but rather are an effect of local management strategies on pipe survival. For example DG1 was often referred to by the interviewed experts (Table 13, Appendix A) as the “problem child”. Lacking corrosion protection and aggravating exposure to electrical currents from households grounding electric appliances on the water pipelines has led to major *pro-active* replacement campaigns. It could furthermore explain the much steeper survival curve of this material in the investigated data set. From consultation with a local expert from whose utility the data was taken, a substantial fraction of pipes is usually replaced before the end of its technical service life, owing to coordination efforts by different network utilities. For instance, the rehabilitation of the sewer system often requires the removal of the above lying water supply pipes. If a substantial part of the replacement is for reasons other than technical end of life, the consequence are considerably shorter observed lifetimes.

The available data neither indicate the reason nor the condition of the replaced pipes. The available survival data does not allow for the description of aging-induced technical (or structural) service life, as it is managerial replacement which is recorded. This means that the expert prior and the data to be combined by Bayesian inference describe two unalike phenomena: the prior describes technical aging and replacement according to the experts' experience, whereas the data represent the aging and replacement process distorted by managerial replacement strategies.

The easiest way to avoid this discrepancy is to solely use survival data of pipes that were replaced due to technical end-of-life, thereby creating congruent information pools.

Nevertheless, this problem easily develops into a philosophical one. Someone may perhaps anyway distrust experts' capacity of differentiation between observed managerial replacement and the replacement caused by structural aging processes. This would mean that the obtained prior and the recorded data do not contradict each other. But who can be trusted if not the interviewed experts who show themselves very able to give such estimates?

Otherwise, the analyst could try to rectify the recorded managerial replacement with a correction factor, thus making it comparable to the experts prior and suitable to describe technical replacement needs. Additional prior estimates for all pipe groups, however, would be needed; something which is already hard to expect from the current managerial strategy. But how to quantify the impact of former management? A way out seems to be basing models on events more capable of describing technical / structural aging, e.g. failure instead of replacement records. Such records are increasingly available and used in different modeling approaches. Instead of priors for technical replacement, priors about the time of a failure or between failures of different orders could be elicited from experts, if not from other utility data. This does not mean, however, that problems induced by former management such as replacement biases in the data can be completely avoided, so that adaptations to the existing models might be necessary.

Although the obtained prior and data might not entirely describe the same phenomenon resulting in pipe replacement, it is nonetheless important to show the performance of Bayesian inference and the prior impact on the posterior. Even if expert and data can be reconciled, prior data conflicts may arise, for example owing to especially unfavorable conditions inducing faster replacement in the study utility than predicted by the experts.

Table 10: Resulting posterior parameters from Bayesian inference with data from a large water utility in Switzerland. The survival model is a Weibull model with parameters $\theta = (\alpha, \beta)^T$. Parameters are given for the case of inference with all data, and artificial data reductions for a prior either derived from complete pooling or partial pooling. MLE gives the parameters obtained by frequentist MLE from all data.

	Grey cast iron (1930-64)				**Ductile iron (1965-80)**				**Asbestos cement**				**Steel**				**Polyethylene**			
	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)	$\hat{\alpha}$	$\hat{\beta}$	sd($\hat{\alpha}$)	sd($\hat{\beta}$)
a) With complete pooling prior																				
Prior	*3.92*	*81.31*	*0.36*	*2.05*	*3.16*	*59.34*	*0.22*	*1.30*	*3.79*	*86.01*	*0.60*	*3.83*	*3.75*	*73.19*	*0.34*	*1.84*	*4.22*	*78.78*	*0.34*	*0.27*
All data	3.36	62.10	0.13	0.84	3.06	58.10	0.16	1.01	2.93	77.36	0.34	3.04	2.97	59.98	0.12	0.98	1.95	74.94	0.24	4.77
500 pipes	3.50	67.48	0.20	1.20	3.04	58.55	0.16	1.12	-	-	-	-	3.03	63.44	0.16	1.29	-	-	-	-
300 pipes	3.61	70.91	0.24	1.46	3.08	58.76	0.18	1.17	-	-	-	-	3.25	67.01	0.19	1.43	-	-	-	-
150 pipes	3.70	74.66	0.28	1.66	3.09	59.01	0.19	1.17	2.83	77.70	0.33	2.93	3.22	68.91	0.22	1.56	2.16	74.82	0.28	4.43
50 pipes	3.71	78.54	0.31	1.83	3.14	59.26	0.19	1.31	3.32	81.90	0.45	3.53	3.66	71.50	0.29	1.75	3.48	76.11	0.68	4.41
0 pipes	3.90	81.32	0.36	2.11	3.16	59.36	0.22	1.29	3.81	86.03	0.59	3.72	3.74	73.26	0.33	1.84	4.22	78.85	1.02	5.03
b) With partial pooling prior																				
Prior	*3.97*	*81.22*	*0.91*	*12.70*	*3.25*	*59.11*	*0.88*	*7.83*	*3.75*	*86.05*	*0.74*	*24.14*	*3.77*	*73.21*	*0.52*	*2.76*	*4.11*	*74.40*	*0.40*	*1.55*
All data	2.26	48.33	0.17	2.34	3.88	52.61	0.52	2.31	2.94	63.51	0.40	3.77	2.84	56.06	0.12	1.14	2.26	43.31	0.28	7.18
500 pipes	2.38	50.55	0.23	3.09	3.01	56.85	0.45	3.19	-	-	-	-	2.93	58.85	0.16	1.51	-	-	-	-
300 pipes	2.62	53.70	0.29	3.25	3.27	55.64	0.52	3.36	-	-	-	-	3.11	63.43	0.19	1.67	-	-	-	-
150 pipes	3.11	58.12	0.43	3.47	2.87	56.67	0.50	3.73	2.74	64.10	0.36	4.30	3.05	65.62	0.25	2.15	2.46	42.40	0.33	6.90
50 pipes	3.07	60.40	0.49	4.82	2.96	56.80	0.57	4.85	3.50	65.90	0.60	5.77	3.63	69.66	0.40	2.53	3.35	48.47	0.65	10.66
0 pipes	3.95	80.78	0.95	12.82	3.26	59.10	0.86	7.79	3.74	86.21	0.73	23.80	3.79	73.15	0.52	2.74	4.08	73.84	1.17	25.87
MLE	2.07	45.54	0.20	3.14	5.48	47.45	0.74	1.82	2.46	60.90	0.85	6.21	2.46	48.77	0.13	1.58	1.81	57.38	0.18	2.37

3.3.5 Bayesian inference

Bayesian inference was performed using the pooled results from expert elicitation as prior and the pipe replacement data. This was also done for different amounts of data. The resulting posterior mean parameters and standard deviations are given in Table 10. Important remark: Following the discussion in section 3.3.4, the survival functions from Bayesian inference described below are not valid for the prediction of technical rehabilitation demand in the studied utility, as the data is about actual replacement that includes replacement for reasons other than technical aging. Nevertheless, the discussion regarding the suitability of the MLE and Bayesian approaches for scarce data situations is valid, as are the observations regarding different prior aggregations.

The most important observation from Table 10 is that in contrast to MLE, reasonably certain parameter estimates could be determined even for small numbers of pipes. This applies especially to pipe groups with few records where MLE returns parameter estimates with high uncertainty (e.g. AC and PE with 150 or less records, or 50 pipes for GI3 and DI1). There, the posterior distributions are more informative than any of the obtainable distributions from MLE alone. For larger data sets however, MLE does lead to reliable parameter estimates which are closer to the mean parameters obtained from MLE of the full data set, making Bayesian combination of data and expert knowledge unnecessary.

Furthermore, as typical for Bayesian inference, the posterior mean values lie between the prior and the MLE (utility data) mean. Also, the standard deviations of the inferred posteriors are smaller than the prior and MLE standard deviations (except for DI1 and PE, see Table 8), meaning that something could be learned from the data. Analogous to MLE, the uncertainty of the parameter estimates increases with decreasing number of pipes used for inference. The smaller the number of pipes, the more the posterior approximates the expert prior (see Figure 6).

In Table 10 and Figure 6, the influence of the prior distribution on the posterior parameter estimates is clearly visible. The wider, partial pooling prior naturally causes wider posterior distributions than the posterior calculated from the more precise complete pooling prior. It also approximates the mean parameters obtained from MLE more closely than the posterior obtained with the complete pooling prior (compare the vertical lines in Figure 6 representing the means of MLE and posterior). In some cases however, the complete pooling posterior coincidentally performs slightly better (being nearer to the MLE estimate).

Regarding the influence of prior aggregation on the posterior, the effect of prior choice is exemplarily shown for 3rd generation grey cast iron in Figure 6. The posterior obtained from inference with the more uncertain partial pooling posterior is notably closer to the MLE than the posterior obtained from inference with the complete pooling prior (Table 10). This effect is attributable to arising prior-data conflicts resulting from discordant information from the observed data and the prior (Bousquet, 2008, among others). Even though the parameter means ($\hat{\alpha}, \hat{\beta}$) of the complete and partial pooling prior are nearly identical, the larger standard deviations ($sd(\hat{\alpha}), sd(\hat{\beta})$) of the partial pooling prior reduce the conflict with the data. No satisfactory approximation towards the MLE shape parameter ($\hat{\alpha}$) of DI1, however, was achieved with any of the two priors. The conceptually more appropriate partial pooling prior (see 3.3.3) leads to a compromise between the prior and the data (posterior $\hat{\alpha} \approx 3.88$ as opposed to prior 3.25 and MLE 5.48) and the complete pooling to hardly any change (posterior $\hat{\alpha} \approx 3.06$, prior $\approx$ 3.16). Checking for conflicts in the distributions in Figure 7 (dotted and dashed lines) as opposed to the distribution obtained from MLE (shaded), it is visible that partial prior and data distributions only partly overlap with approximately similar densities, but the complete pooling prior and the data widely disagree.

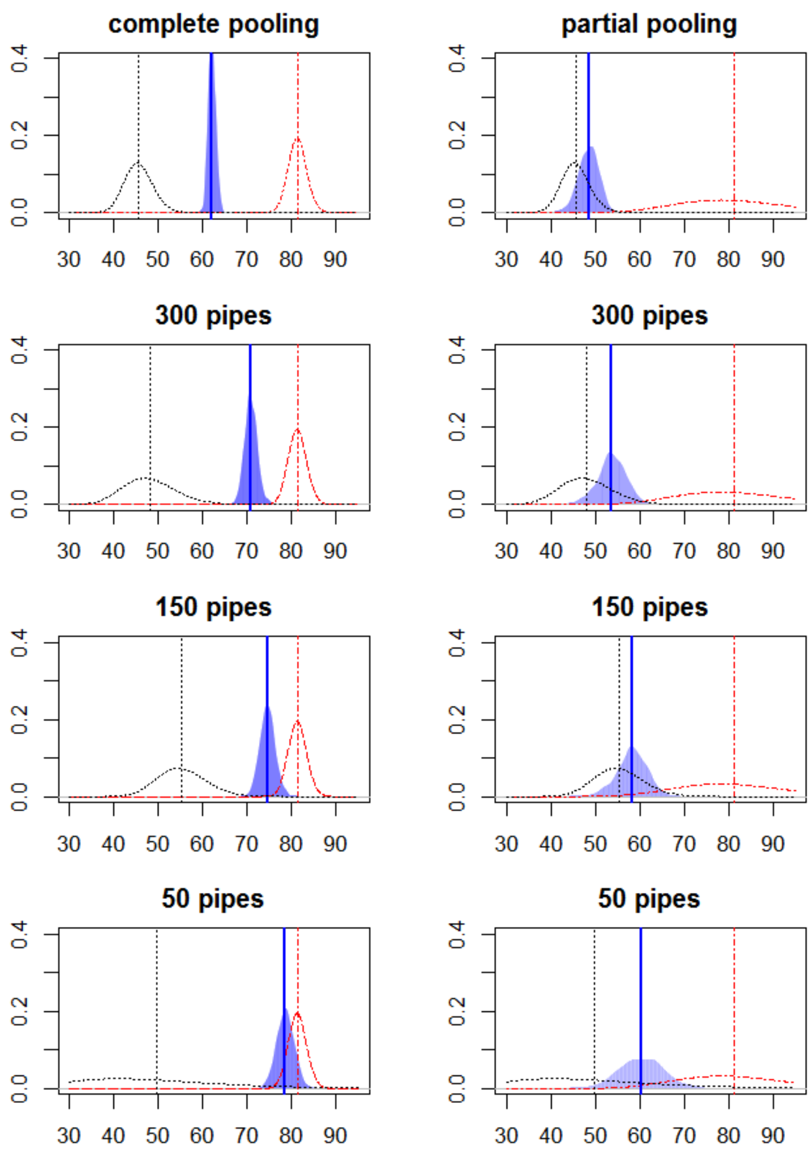


Figure 6: Bayesian inference with a complete (left) and partial (right) pooling prior for GI3. Prior (red dash-dotted), posterior (blue filled), and MLE (black dotted) marginal density distributions of the Weibull scale parameter μ_β are shown. Vertical lines indicate the position of the corresponding means. The top row shows the inference results using all data. Note that for no utility data (0 pipes, not shown), the posterior and prior marginal distributions are coincident.

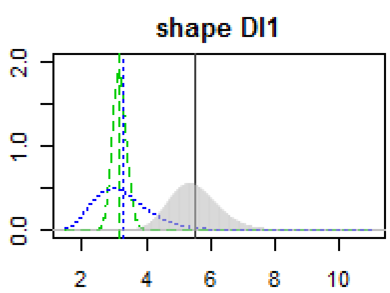


Figure 7: Shape parameter μ_α distributions of the two priors compared with the MLE (grey- shaded, solid) for DI1. Partial pool: blue dotted lines; complete pool: green dashed lines. Vertical lines indicate the position of the corresponding means.

3.4 Summary and Conclusions

3.4.1 Improved service life modeling under scarce data

We suggest a systematic approach to water pipe service life modeling that uses expert information from several utilities as prior which is then updated with local utility data. For this purpose available methods of expert elicitation of an unknown probability distribution were adapted and furthermore extended to imprecise quantile estimation of the pipe survival function. Contrary to currently existing approaches, encouraging experts to state interval estimates for the quantiles, leads to the imprecision or uncertainty of the expert being explicitly included in the analysis. From these statements, a bivariate expert prior for the survival function parameters is inferred, thus overcoming the difficulties confronted in elicitation of multivariate (i.e. correlated) distributions. The resulting expert priors can be aggregated with a linear pooling approach to get an intersubjective prior approximating the state of knowledge across experts and environmental conditions. For both Bayesian and frequentist inference of the parameters of the survival function from utility data, a likelihood function for the commonly encountered left-truncated right-censored pipe network data is derived.

The results from section 3.3.5 testify that the proposed approach improves estimation of the expected service life of water networks under scarce data, leading to the ability to identify parameters where otherwise not possible or to derive more informative estimates than with using MLE alone. This is a key improvement for more effective rehabilitation planning and water distribution network management.

3.4.2 Expert elicitation and prior aggregation

Priors for the parameters of the survival function characterizing technical service life of five pipe groups were obtained from interviews with eight water utility experts, ordered by perceived durability, most durable to least: asbestos cement, grey cast iron (1930-1965), polyethylene, steel, and ductile cast iron (1965-1980). This durability order is in agreement with literature estimates from Germany and Austria; the specific lifetime may vary depending upon local conditions.

An important aspect of our approach is the incorporation of the between-experts-variance into the aggregated prior distribution from individually fitted expert distributions. It is shown that not only the uncertainty of the single-expert, but also the deviation between experts is a valuable source of knowledge in itself, as it covers the different network conditions from various utilities. The resulting partial pooling prior of the experts' distributions leads to an intersubjective prior that covers these different water distribution network conditions as well as the different opinions of the experts. Only if the aim was a specific prior for a single utility and if more than one expert representing this utility was available, the complete pooling prior would be more appropriate because differences between experts are interpreted as measurement errors.

3.4.3 Frequentist and Bayesian inference

The results reviewed in section 3.3 suggest that Bayesian inference of survival function parameters by considering expert knowledge is a suitable approach to bridge the scarce data situation encountered in many water utilities. Frequentist estimation remains the less demanding approach if sufficient data is available (roughly more than 150 pipes with at least 5 years of data for this utility). To avoid prior-data conflicts, the validity of prior and posterior distributions for the locally encountered conditions can be discussed with a local expert.

3.4.4 Ambiguity of model selection

The problem of ambiguity arising in model selection is addressed by fitting three standard parametric survival models to the experts' data and choosing the best fitting for all experts. The Weibull model overall provided the best fit to experts statements, but did not prove the best choice for all pipe groups and experts.

Following the discussion in section 3.3.4, the use of pipe replacement data for prediction of a network's technical rehabilitation demand needs to be revised. Replacement data alone are not suitable to predict structural aging only, but also reflect managerial decisions. In order to adopt more efficient rehabilitation approaches, a predictive model needs to be able to quantitatively describe these two factors separately. As the presented model might be prone to systematic biases induced by these limitations, it is not meaningful to discuss its predictive capability. The model's goodness of fit is nevertheless demonstrated by the close overlay of the non-parametric and parametric model shown in section 3.3.2.

3.4.5 Consideration of uncertainty

Tackling uncertainty on different levels and during different steps is useful to get an overall assessment of uncertainty the assessed quantities. This involves the use of examples and visual support tools to help the experts correctly express their belief in a quantitative manner. Cognitive biases can be reduced through training and control during elicitation, besides double-checking and repetition of the stated values. Influence of model selection can be elucidated by fitting several models and choosing the one that minimizes the deviation between measurements and modeled data. Lack of fit to the stated expert quantiles is made explicit by measurement of the error in the parameters.

3.5 Acknowledgements

This research was funded within the National Framework Program 61 on Sustainable Water Management by the Swiss National Science Foundation (SNF). We thank the eight anonymous experts, the Swiss Gas and Water Association (SVGW), as well as Sébastien Apothéloz and Eauservice Lausanne (the large water utility providing replacement data), for their cooperation and contribution in this study. Special thanks owe to Adelaida Arreaza, Christoph Egger, and Simon Rinderknecht for their important contributions in preparing and performing the interviews with the experts, as well as colleagues for the valuable discussions.

3.6 Appendices

Appendix A- Expert elicitation

Table 11: Description of pipe groups based on common differentiation criteria.

pipe group	abbre-viation	differentiation criteria	chosen for elicitation and underlying rationale
asbestos cement incl. Eternit	**AC**	years 1930-1985	Yes, although no longer produced. Has proved to be very resistant if appropriately installed and is still common in Swiss water networks, especially in smaller communities, under the name of "Eternit".
1st generation grey cast iron	**GI1**	from horizontal sand molds, corrosion-resistant, varying wall thickness; before 1880	No, because close to no occurrence in today's Swiss networks.
2nd generation grey cast iron	**GI2**	vertically cast, more corrosion resistance, ca. 1880-1930	No, mostly in bigger cities only; building of networks in smaller communities rather later.
3rd generation grey cast iron	**GI3**	centrifugal casting, susceptible to corrosion, 1930-1965	Yes, slowly being replaced, but still common.
1st generation ductile cast iron	**DI1**	centrifugal casting, lacking external corrosion protection, high tensile strength, 1964-1980	Yes, slowly being replaced, but still common.
2nd generation ductile cast iron	**DI2**	similar to DI1, but improved external corrosion protection, after 1980	No, because probably only little replacement up to today and time constraints. Interesting in hindsight because of the rather large proportion in today's networks.
1st generation steel	**ST1**	welded or seamless, lacking external corrosion protection, before 1930	Yes, but without differentiation into generations because occurrence of this material varied in the networks the experts were familiar with and they were not confident in making further differentiations.
2nd generation steel	**ST2**	insufficient external corrosion protection, ca. 1930 - 1980	
3rd generation steel	**ST3**	enhanced external corrosion protection, after 1980	
PVC polyvinylchloride	**PVC**	approx. 1930-1990	No, only rarely used in Switzerland's water supply systems.
1st generation polyethylene	**PE1**	PE-LD; PE-HD; and PE 63; before 1980	Yes, but also without differentiation because it is a rather new application in Switzerland and because of the high expected service life. Not yet many replacements if installed correctly (leading to a lack of experience of experts with this material).
2nd generation polyethylene	**PE2**	PE 80 and PE 100; after 1980	
3rd generation polyethylene	**PE3**	PE-X, cross-linked PE, high toe-crack resistance; after 1992	

Table 12: Stated quantile values from expert elicitation, l = lower value, u = upper value. If only one value is given, the expert stated precise estimates. The first column indicates material and quantile.

	E1		**E2**		**E3**		**E4**		**E5**		**E6**		**E7**		**E8**	
group	l	u	l	u	l	u	l	u	l	u	l	u	l	u	l	u
GI3																
0.05	30		20	40	30		30	40	20	30	40		50		60	70
0.25	40		40	60	65		40	50	30	50	55		70		60	90
0.5	60		70	90	80		60	70	60	80	80		90		90	110
0.75	80		80	100	85		70	85	70	90	90		100		100	130
0.95	100		90	110	110		85	90	80	110	100		120		130	140
DI1																
0.05	20		10	15	25	30	10	20	20	30	20	30	20		30	40
0.25	30		20	30	40	55	20	30	30	40	40	50	45		40	60
0.5	45		40	60	60	70	45	55	40	60	50	60	55		60	70
0.75	65		60	70	70	80	55	60	60	70	60	70	70		80	90
0.95	100		70	80	80	85	60	70	80	90	70	80	80		90	100
AC																
0.05	30	50	40		40	60	-		20		-		20	30	-	
0.25	60	80	80		75	85	-		35		-		40	60	-	
0.5	80	100	100		90	110	-		50		-		55	65	-	
0.75	100	120	110		115	135	-		65		-		65	75	-	
0.95	120	150	120		140	150	-		80		-		75	90	-	
ST																
0.05	-		-		-		20		-		-		30	40	30	40
0.25	-		-		-		41		-		-		45	55	40	60
0.5	-		-		-		42		-		-		60	70	60	70
0.75	-		-		-		43		-		-		70	90	80	90
0.95	-		-		-		45		-		-		90	110	90	100
PE																
0.05	0	50	40		-		30		15		50	60	-		65	
0.25	50	150	60		-		40		25		60	80	-		80	
0.5	100	150	80		-		50		35		80	100	-		100	
0.75	100	150	95		-		60		55		100	110	-		120	
0.95	100	200	110		-		70		70		100	150	-		140	

Table 13: Description of experts and locally specific influence factors of network deterioration

Expert	Position / qualification	Mentioned influence factors
E1	Planning & construction engineer; head of local engineering company servicing several small water suppliers in the Zürcher Oberland	- DI1 is "problem child" - GI3 has mostly been removed (only about 5-10 % of today's network) - bedding - AC: connections deteriorate faster than pipes
E2	Planning & construction engineer of the same local engineering company as expert 1; but servicing different communities	- GI3 ca. 30 % of current network - settling and corrosion problems with GI3 - bedding - electrical grounding of house installations on DC1 pipes - PE more or less 20-30 % of network, some problems with earlier PE pipes - AC: connections deteriorate faster than pipes, mostly used for large diameter transport pipes
E3	Operation & maintenance engineer; head of distribution network department of the public water supply for a medium size city in NE Switzerland	- close to no PE and ST in current network - electrical grounding of house installations on DC1 pipes until late 90s - GI3 approx. 18 % of network - DI1 approx. 20 % of network, usually 45 years of service life assumed - AC: laying depth is rather deep, mostly used for larger diameter transport pipes - favorable soil conditions
E4	Project and construction manager; head of distribution network department of a private water supply company serving a medium size city in central Switzerland	- laying depth of AC is rather deep - mechanical impacts by traffic - DC1 problems with bedding (wood used as support under the pipe) - many different pressure zones - because of large financial losses caused by GI3 failures, this material is replaced earlier based on risk considerations - ST was massively replaced in the 90s because of failures probably caused by inappropriate bedding
E5	Operation & maintenance engineer of a consortium of small municipal water suppliers in central Switzerland	- strongly varying soil conditions from rugged rocks over river gravel to heavy clay and aggressive moor soils - problems with quality of PE installation, especially welding has been an issue; rather early use of PE; PE only used if conditions do not allow for metal pipes (mostly soil) - rather soft water (12-13 °fH eq. to 1,2-1,3 mmol/L) - tank traffic is problematic for GI3 - overall many different pressure zones
E6	Network utility engineer from a public water supplier of a small city in NW Switzerland	- usually assume fixed service life of 60 years for pipes - strongly favoring PE for replacements; rather high percentage of the network are PE pipes (> 20 %) - problems with both bedding (timber, wooden support) and electrical grounding of house installations on DC1 pipes
E7	Facility manager and engineer of a consortium of small water suppliers in NW Switzerland	- strongly favoring cast iron, PE only for household connections - difficult ground because of soil variations from strongly settling to peaty soils - problems with bedding in 60's and 70's when most of their pipes were installed - have experienced many failures in both GI3 and DI1 pipes
E8	Head of distribution network department of the public water supply of a larger city in NE Switzerland	- assume 100-120 years of service life for GI3 - systematic defects in pipes built shortly after Second World War - earlier PE types expected to have much shorter duration than newer PE types; longer service life assumptions supported by a recent material study of DVGW - heterogenic soil and ground properties - electrical grounding of house installations on DC1 pipes

Detailed description of the elicitation procedure

The interviews were attended by at least three persons: the expert, the interviewer and the analyst. If more people were attending the interviews (e.g. assistants), they were not allowed to actively participate or alter the elicitation procedure. The interviewer and analyst completely abstained from any kind of coaching regarding the order of magnitude of the answers.

Setting the scene: At the beginning, the interviewer repeated the purpose of the study, and explained the way she would proceed during the interview. The aim of the elicitation was clearly stated. Then, the quantity to be elicited was clarified, and the five pipe groups were presented. It was reaffirmed that the named pipe groups actually occurred in the expert's domain. The four most familiar (in one case three) were then selected for elicitation. After this was done, the expert was asked several questions regarding his expertise and familiarity with probability.

If the expert had not read or understood the pre-elicitation information, there was time to go through it in detail.

Focusing: Then, the expert was requested to name the most important influencing factors for pipe aging in his area (see Table 13). This was not only done to learn about special circumstances of the different localities, but also to make him concentrate on the upcoming task. Then, he was motivated by stating that his knowledge was an important source of information for the estimation of pipe service life. It was made explicit that he was not expected to be all-knowing, but that his expertise was crucial for the study. This was to encourage him to answer as best according to his knowledge while adequately stating his uncertainty.

Training: Before elicitation of the quantities of interest started, an elicitation round identical to the assessment to come of pipe service life was done to train the expert. To avoid anchoring, the survival of women born in Switzerland in the year 1940 as provided by the Swiss Federal Statistics Office (Cordazzo, 2006) was used. It was used to familiarize the expert with the elicitation procedure and to avoid possible misunderstandings. It also helped to cross-check for calibration (which was clearly not possible for pipes because of a lack of factual measurement data for the experts utility). This procedure allowed the interviewers to point out features to keep in mind which can lead to biases that might well arise during elicitation of pipe survival.

Elicitation of pipe groups: After this training, the pipe group being addressed was again specified. The expert was asked about his experience with this group in his area of responsibility. Then, the quantiles were elicited in a first round following the sequence:

Define the overall interval defined by the age at which most pipes are expected to have been replaced (95 %) and at which close to all pipes of a group / cohort are still in service (5 %).

The age at which half (50 %) of the pipes are replaced and half remain in service is the median. This is the third quantity elicited.

Finally, the values for three-quarters (75 %) and a quarter (25 %) are assessed.

At the same time, the addressed quantities were roughly visualized using 100 paper clips (representing 100 % of the pipe group) and a paper sheet with the time bar. Experts were requested to disregard replacement because of initial laying failures (e.g. within the first year after laying), as well as replacements following managerial or economic considerations, such as coordinative ground works with other infrastructure providers. This helped to focus on technical aging and not on effects of different management decisions.

Secondly, the quantiles for 75 %, 50 %, and 25 % were re-elicited using bets, adjusting the stated ages until the expert was indifferent between the bets. This technique was used by the interviewer to confirm the statements by making the experts think differently about the quantities. After the bets, she read out the documented values and individually confirmed them with the expert. These checks and repetitions were done to ensure that experts' statements are reliable, consistent and correctly documented. At the end, the experts were asked to provide a qualitative description of the imaginary density curve, if possible. The description should reveal whether it could be assumed unimodal.

This elicitation procedure was repeated for the selected pipe groups. During the interview, the experts were asked to assess half of the pipe groups stating imprecise estimates and the other half using precise estimates.

Feedback: At the end of the interview, the experts had to assess the difficulty of the interview, how realistic they think their stated answers are, and if they preferred stating point estimates or intervals.

Table 14: Estimates of pipe survival in the literature as reported from Austria and Germany

	Roscher et. al (2005)			**Fuchs (2001)***			**Trujillo Alvarez (1995)***		
	0 %	**50 %**	**90 %**	**0 %**	**50 %**	**90 %**	**0 %**	**50 %**	**90 %**
AC	-	-	-	20-50	60-90	80-110	5-80	50-90	60-110
GI1	60-70	65-90	80-110	-	-	-			
GI2				30-80	100-160	130-190	5-70	40-101	80-150
GI3	40-60	65-90	80-100	10-30	50-90	70-110			
DI1	20-30	45-65	70-100	5-20	40-70	70-90	10-60	30-90	50-120
DI2	40-50	75-90	100-130	80-120	100-140	120-160	6-100	50-140	90-165
ST1				-	-	-			
ST2	40-50	60-80	80-110	-	-	-	4-60	25-100	55-120
ST3				100	120	140			
PVC	10-30	30-50	50-70	10-30	40-85				
PE1	15-30	50-75	50-70	-	-	-	5-60	40-80	50-100
PE2	40-50	75-90	100-130	-	-	-			
PE3	40-50	75-90	100-130	20-30	50-70	80-90	-	-	-

* Originally, intervals for an optimistic and pessimistic estimate were given which were herein merged together. Thus, the upper bound of the stated interval is the upper bound of the optimistic estimate; the lower bound equals the lower bound of the pessimistic estimate.

Appendix B- Parametric model fit and parameter uncertainty

Table 15: Goodness of fit of Weibull, lognormal, and gamma distribution from non-linear least squares regression over the elicited quantiles. Bold numbers indicate the model which minimizes the residual sum of squares (RSS) and residual standard error (RSE), dependent on available degrees of freedom (doF).

			Weibull		lognormal		gamma	
group	**expert**	**doF**	RSS	RSE	RSS	RSE	RSS	RSE
GI3	E1	3	**59.6**	4.458	95.8	5.65	61.9	4.542
GI3	E2	8	**1380**	13.14	1930	15.53	1740	14.74
GI3	E3	3	**140**	6.837	342	10.68	272	9.515
GI3	E4	8	**384**	6.932	526	8.112	467	7.641
GI3	E5	8	**1370**	13.1	1780	14.91	1610	14.18
GI3	E6	3	**89.5**	5.463	240	8.937	190	7.963
GI3	E7	3	**14.1**	2.169	100	5.781	65.2	4.663
GI3	E8	8	1420	13.32	1390	13.18	**1360**	13.03
DI1	E1	3	62.1	4.548	**11.6**	1.964	15.4	2.267
DI1	E2	8	**650**	9.015	1030	11.37	850	10.31
DI1	E3	8	**375**	6.851	698	9.338	595	8.626
DI1	E4	8	**497**	7.885	780	9.873	664	9.108
DI1	E5	8	452	7.515	462	7.6	**430**	7.328
DI1	E6	8	**292**	6.046	482	7.765	419	7.236
DI1	E7	3	**61.5**	4.526	217	8.506	159	7.285
DI1	E8	8	**492**	7.846	668	9.138	590	8.591
AC	E1	8	**1270**	12.58	1680	14.48	1490	13.67
AC	E2	3	**277**	9.61	694	15.21	582	13.93
AC	E3	8	**754**	9.705	1220	12.35	1030	11.35
AC	E5	3	**10.8**	1.899	96.3	5.665	53.1	4.205
AC	E7	8	**551**	8.302	809	10.06	723	9.506
ST	E7	8	**577**	8.491	620	8.805	577	8.489
ST	E8	8	**492**	7.846	668	9.138	590	8.591
PE	E2	3	**22**	2.709	162	7.343	108	5.998
PE	E4	3	**6.54**	1.476	23.2	2.782	13.5	2.125
PE	E5	3	**32.5**	3.293	68.8	4.789	3.711	41.3
PE	E6	8	**1820**	15.09	1860	15.26	**1820**	15.07
PE	E8	3	67.3	4.736	63.3	4.593	**42.9**	3.782

Estimation of the variance-covariance matrix from the Fisher information matrix

The inverse expected Fisher information matrix is essentially equivalent to the negative of the Hessian matrix of the obtained estimate, i.e. the second order partial derivatives. In the case of a two-parametric model the Hessian is a 2,2-square matrix calculated from:

$$H(\ln L) = \frac{\delta\theta_2{}^2\delta\theta_1}{\delta\theta_1{}^2}\ln l(\boldsymbol{\theta}) = \begin{bmatrix} \frac{\delta^2 lnL}{\delta\theta_1^2} & \frac{\delta^2 lnL}{\delta\theta_1\delta\theta_2} \\ \frac{\delta^2 lnL}{\delta\theta_2\delta\theta_1} & \frac{\delta^2 lnL}{\delta\theta_2^2} \end{bmatrix} = \begin{bmatrix} \delta_{\theta_1\theta_1} & \delta_{\theta_1\theta_2} \\ \delta_{\theta_2\theta_1} & \delta_{\theta_2\theta_2} \end{bmatrix} \tag{12}$$

where θ_1,θ_2 are the parameters of the parametric model. The Hessian (matrix) describes the local curvature of the logL function. This, and the reasoning behind it are well described in Harrell (2001, pp. 180-183).

Description of the extended Nelson-Aalen estimator

A Nelson-Aalen estimator adapted for left-truncated and right-censored data (also referred to as *extended Nelson estimator)* as described in Pan and Chappell (1998) and applied to pipe survival in Carrión, Solano et al. (2010)) is used.

$\hat{H}_a(t) = \sum_{a \le t_i \le t} \left[\frac{d_i}{Y_i}\right], t \ge a,$ (Klein and Moeschberger, 2003)

t … age at replacement or censoring (=end of study)

a_i… age at entering the study

d_i…number of events(=deaths / replacements) at time t_i

Y_i … number of individuals entering the study before t_i and for which a< $t_i \le$ t

the survival can then be estimated with:

$$\hat{S}_a(t) = e^{-\hat{H}_a(t)}$$

With this layout, over-estimation through right-censoring and underestimation caused by left-truncation is reduced / avoided. According to (Klein and Moeschberger, 2003), this estimator can be interpreted as an estimator of the probability of survival beyond time t conditional on the smallest of the entry times: Pr[T > t|T > a]. T: time to event.

Calculation of point-wise confidence intervals of nonparametric fit

(Klein and Moeschberger, 2003) from S.105ff:

log-transformed confidence intervals are chosen because of better reported performance in small samples

$$\left[\hat{S}_{t_0}^{1/\theta}, \hat{S}_{t_0}^{\theta}\right];\ \theta = e^{\left\{\frac{Z_{1-\alpha/2}\cdot\sigma_S(t_0)}{\ln[S(t_0)]}\right\}}$$

with $\boldsymbol{\sigma_S^2(t)} = \frac{\hat{V}[\hat{S}(t)]}{\hat{S}^2(t)}$

The resulting confidence interval is not necessarily symmetric about the estimate of the survival function.

4 Extension of pipe failure models to consider the absence of data from replaced pipes

Andreas Scheidegger[a)], Lisa Scholten[a)], Max Maurer[a,b)], Peter Reichert[a)]

[a)] Eawag, Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, 8600 Duebendorf, Switzerland

[b)] Institute of Environmental Engineering, ETH Zurich, 8093 Zürich, Switzerland

Water Research (2013), 47, 3696-3705. http://dx.doi.org/10.1016/j.watres.2013.04.017

Abstract

Predictions of the expected number of failures of water distribution network pipes are important to develop an optimal management strategy. A number of probabilistic pipe failure models have been proposed in the literature for this purpose. They have to be calibrated on failure records. However, common data management practices mean that replaced pipes are often absent from available data sets. This leads to a 'survival selection bias', as pipes with frequent failures are more likely to be absent from the data. To address this problem, we propose a formal statistical approach to extend the likelihood function of a pipe failure model by a replacement model. Frequentist maximum likelihood estimation or Bayesian inference can then be applied for parameter estimation. This approach is general and is not limited to a particular failure or replacement model. We implemented this approach with a Weibull-exponential failure model and a simple constant probability replacement model. Based on this distribution assumptions, we illustrated our concept with two examples. First, we used simulated data to show how replacement causes a 'survival selection bias' and how to successfully correct for it. A second example with real data illustrates how a model can be extended to consider covariables.

Keywords

Pipe failure model, replacement model, likelihood, Bayesian inference, survival selection bias

4.1 Introduction

The optimal management strategy for water distribution networks balances issues of water safety, reliability, quality, and quantity, while exploiting the full extent of the useful life of the pipes to achieve economic efficiency (Kleiner and Rajani, 2001). Pipe failure models are one of the key tools to support this management process.

We distinguish between two major applications of pipe failure models: (i) The failure probabilities of the individual pipes are needed for the mid-term maintenance and replacement strategies of the pipe network (ii) For long-term planning, the expected number of failures in the entire system is of interest, but not the specific cause of the failures. It is therefore sufficient to model all deterioration processes lumped together as a function of age, so that less detailed data are required. Applications (i) and (ii) do not require fundamentally different model structures, because models for (ii) can typically be extended to fulfill the needs of (i) by incorporating pipe properties such as material, diameter, etc. to improve pipe-specific predictions.

The model should be calibrated on the basis of failure records of the local system because of differences in the influence factors that are not modeled (e.g. soil properties). Correct calibration can become challenging because the available data typically show some or all of the following properties (see also Figure 8):

- Right censored observations (Figure8-i): For every pipe in service a right censored observation is available: the time since the last failure or construction until the time of observation. This provides important information, and pipes without recorded failures until the end of the observation period must not be excluded from the calibration process. This issue is considered in the calibration procedures of many time-based failures models (Carrión et al., 2010; Eisenbeis et al., 1999; Gustafson and Clancy, 1999; Mailhot et al., 2000). For models formulated as a counting process (e.g. Economou et al., 2009; Kleiner and Rajani, 2010; Watson et al., 2004), right censoring is not relevant, because the probability of a certain number of failures within a time interval is modeled instead of the time between the failures.

- Left truncation (Figure 8-ii): Left truncation occurs if a pipe was installed before failures were systematically recorded by the utility. As a consequence, it is not known how many failures occurred before the recording period. Only few models (Carrión et al., 2010; Le Gat, 2009; Mailhot et al., 2000) explicitly consider left truncation.

- Absence of replaced pipe data (Figure 8-iii): Frequently, replaced pipes are deleted from the database together with the corresponding pipe failure data because the database was established with the objective of reflecting the current state. This leads to a "selective survival bias" (Renaud et al., 2012), due to the fact that pipes with poor failure histories will be underrepresented in the data set. Hence, ignoring this in the parameter estimation causes systematic errors in the predictions which cannot be reduced merely by increasing the amount of data (Scheidegger et al., 2011).

The intuitive idea to consider the survival selection bias is to assess how likely it is that a pipe similar to the observed has been replaced in the past and correct the likelihood function accordingly. This requires the integration of a *replacement model* that characterizes the probability that a pipe was not replaced, i.e. the chance that a pipe is still in service. Generally, this can be a function of the condition, age and number of failures a pipe has already experienced. The parameters of the replacement model are then estimated jointly with those of the failure model.

To the best of our knowledge only the LEYP model (linear extension of the Yule process) developed in the dissertation of (Le Gat, 2009) attempts to tackle the selective survival bias. Le Gat (2009) specified the probability of a pipe is not being replaced if a failure occurs as a function of pipe age. The chosen double exponential form allows an analytical evaluation of the likelihood of the LEYP model. In the form presented, however, this approach is difficult to generalize and to transfer to other failure and replacement models. For example, a replacement decision might not depend on age but on the number of previous failures.

In this paper we propose a general framework to derive the likelihood function of a pipe failure model combined with any kind of replacement model to enable unbiased parameter estimation from data sets without historical records. The likelihood function derived in section 4.2 has a frequentist interpretation so that the parameters can either be estimated according to the maximum likelihood principle or by Bayesian inference. The latter is favorable in two common cases: (i) Small utilities often have very limited data, either because they have been recording failures only for a short time or simply because they have small networks. However, they typically have dedicated experts with sound practical experience beyond the information in archives. In this situation, carefully elicited expert knowledge can improve the model performance (Scholten et al., 2013). (ii) The parameters of the replacement model can correlate strongly with those of the failure model and may therefore lead to identifiability problems in a frequentist setting. In a Bayesian framework, this can be circumvented by an informative prior distribution.

The remainder of this paper is structured as follows. In section 4.2 we first introduce a universal notation for pipe failure models and then derive the likelihood function for completely and partially observed pipes. On this basis, we illustrate how a failure model can be extended with a replacement model in general. Furthermore, the predictive distributions for the number of failures for pipes with and without failure record are presented. As an example, the equations are derived explicitly for a Weibull-exponential model in section 4.3. In the following section, this model is used for two application examples: the first is based on artificial data to highlight the importance of the replacement model. The second illustrates (with real data) how individual pipe properties can be considered. Finally, we discuss the strengths and weaknesses of our approach and point out directions for further research.

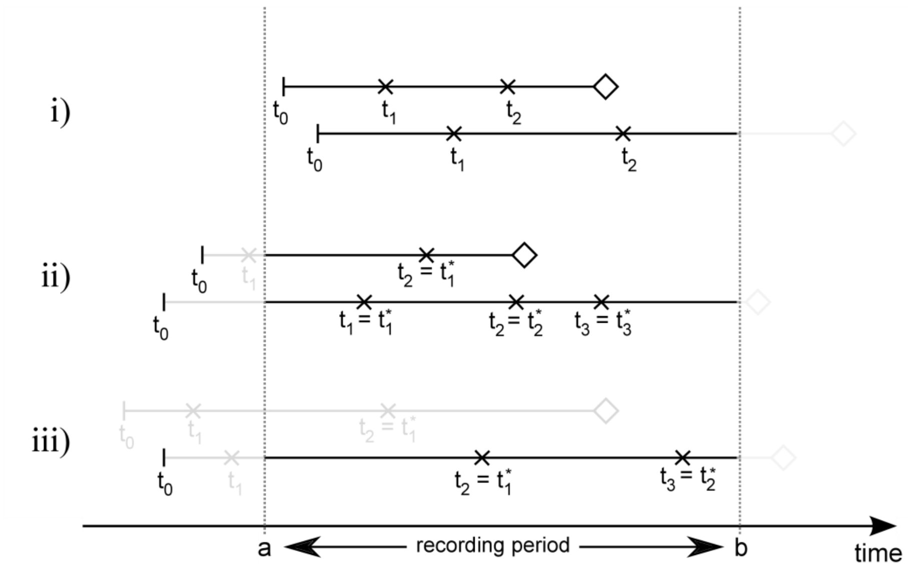


Figure 8 Three scenarios with different data availabilities. The available information is shown in black, the unavailable information in gray. a marks the beginning and b the end of the recording period, × a failure, ◇the replacement, $\boldsymbol{t_0}$ is the time of construction, $\boldsymbol{t_i}$ the time of the *i*th failure, and $\boldsymbol{t_i^*}$ the time of the ith recorded failure. i) All failures of the pipes are recorded and data of replaced pipes remain in the data set; ii) Failures before a are not recorded, the number of failures per pipe is unknown; iii) The total number of failures per pipe is unknown and data of replaced pipes is unavailable.

4.2 Methods

4.2.1 Pipe failure model

As long as a pipe is in operation there is a chance of a failure event. We define a failure as an observable event that requires immediate measures (e.g. a break). Other definitions are possible, depending on the available records. It is assumed that failures are repaired immediately without replacing the pipe.

For a single pipe, the point in time of the *i*th failure is denoted by t_i while t_0 stands for the time of construction. The time when the *i*th failure occurs is random and therefore described by a probability density function $p_i(t|t_0, \ldots, t_{i-1,}\boldsymbol{\theta})$or a survival probability $Prob(t_i > t|t_0, \ldots, t_{i-1}, \boldsymbol{\theta}) = S_i(t|t_0, \ldots, t_{i-1}, \boldsymbol{\theta})$. Obviously, the $t_i, i \geq 0$ are not independent as the *i*th failure cannot occur before the $(i-1)$th failure. The vector $\boldsymbol{\theta}$ represents the parameters of all distributions.

This formulation enables us to express different standard models with the same notation. Models based on a counting process can be written equivalently as time-based models. For example for a homogeneous Poisson process, we would define $p_i(t|t_0, \ldots, t_{i-1,}\boldsymbol{\theta}) = p(t - t_{i-1}|\boldsymbol{\theta}) = \lambda e^{-\lambda(t-t_{i-1})}$ for $i > 0$, i.e. the time differences between two failures are all exponentially distributed with the same rate λ.

To statistically estimate the parameters $\boldsymbol{\theta}$ and for failure predictions, a likelihood function is required. The likelihood is the joint probability (density) of the observed n_k failures at times $\mathrm{T}_k = \{t_{k,i}: \mathrm{i} = 0, \ldots, n_k\}$ for all pipes $k = 1, \ldots, K$ given the model and the parameters.

We define *a* as the time at the beginning of the recording period and *b* as the end of the recording period. All failures are recorded within this period. The likelihood function for a pipe failure model for two data collection schemes is derived below (compare Figure 8): i) the complete life of the pipe lies within the recording period, and ii) the pipe was built before recording started.

All the following equations apply to a single pipe unless otherwise stated. For the sake of simpler notation, the pipe index k is omitted in equations that refer to a single pipe.

Likelihood for completely observed pipes

If $a \leq t_0$, the recording period covers the complete life of the pipe. For this situation the likelihood of *n* failures at times $\mathrm{T} = \{t_i: \mathrm{i} = 0, \ldots, n\}$ for one pipe is formulated as

$$p(\mathrm{T}, \mathrm{n}|\mathrm{b}, \boldsymbol{\theta}) = \left[\prod_{i=1}^{n} p_i\,(t_i|t_0, \ldots, t_{i-1}, \boldsymbol{\theta})\right]\, S_{n+1}(b|t_0, \ldots, t_n, \boldsymbol{\theta}) \tag{13}$$

where $\boldsymbol{\theta}$ represents the parameters of the distributions. The factor $S_{n+1}(b|t_0, \ldots, t_n, \boldsymbol{\theta})$ accounts for the fact that there is always a right censored observation available: the time from the last failure (or from construction) until the end of the observation period or the replacement of the pipe.

Likelihood for partly observed pipes

If a pipe was built before the observation period began $(t_0 \leq a)$ it is not known how many (if any) failures have occurred before *a*.

The likelihood proposed by Mailhot et al. (2000) accounts for this. In the following a distinction must be made between t_i^*, the point in time of the ith *recorded* failure and t_i, the time of the ith failure which is not necessarily equal to t_i^*. The n recorded failures are summarized as $\mathrm{T}^* = \{t_i^*: i = 0, \ldots, n\}$. Additionally the time of construction t_0 is assumed to be known. For convenient notation we define $t_i^* := t_0$. Note that $p_i(t|\boldsymbol{\theta}_i)$ still stands for the density of the time of the ith (observed or unobserved) failure.

Mailhot et al. (2000) first derived the joint distribution of the number of non-recorded failures *m* and the *n* recorded failures at T^*. Adapted to our notation and slightly generalized, this is written as

$$\begin{aligned} p(\mathrm{T}^*, m, n|a, b, \boldsymbol{\theta}) = \int_{t_0}^{a}\int_{t_1}^{a} \ldots \int_{t_{(m-1)}}^{a} p_1(t_1|t_0)\, p_2(t_2|t_{0,}t_1) \ldots\, p_m(t_m|t_0, \ldots, t_{m-1}) \,\cdot \\ p_{m+1}(t_1^*|t_0, \ldots, t_m)\, p_{m+2}(t_2^*|t_0, \ldots, t_m, t_1^*) \ldots p_{m+n}(t_n^*|t_0, \ldots, t_m, t_1^*, \ldots, t_{n-1}^*) \,\cdot \\ S_{m+n+1}(b|t_0, \ldots, t_m, t_1^*, \ldots, t_n^*) dt_m \ldots dt_2 dt_1 \end{aligned} \tag{14}$$

for $m > 0$. For no non-recorded failures, $m = 0$, the density $p(\mathrm{T}^*, m = 0, n|a, b, \boldsymbol{\theta})$ takes the form of (13).

The likelihood for a single pipe is then obtained by summing (14) over *m*:

$$p(\mathrm{T}^*, n|a, b, \boldsymbol{\theta}) = \sum_{m=0}^{\infty} p(\mathrm{T}^*, m, n|a, b, \boldsymbol{\theta}) \tag{15}$$

4.2.2 Replacement model

The replacement model has to express the probability of the event 'pipe has not been replaced up to time *b*' (abrv. 'not rep.') given its failure history, $Prob('not\ rep.'\,|\mathrm{T}^*, n, a, b, \boldsymbol{\theta})$.

It is usually more convenient to formulate the replacement model first conditioned on the number of non-recorded failures, i.e. $Prob('not\ rep.'\,|\mathrm{T}^*, n, a, b, \boldsymbol{\theta})$. The unconditional replacement model is then derived as

$$Prob('not\ rep.'\,|\mathrm{T}^*, n, a, b, \boldsymbol{\theta}) = \sum_{m=0}^{\infty} Prob('not\ rep.'\,|\mathrm{T}^*, n, m, a, b, \boldsymbol{\theta}) Prob(m|a, \boldsymbol{\theta})$$

where the probability of *m* failures before *a* is given by

$$Prob(m|a, \boldsymbol{\theta}) = \int_{t_0}^{a} \int_{t_1}^{a} \dots \int_{t_{(m-1)}}^{a} \left[\prod_{i=1}^{m} p_i\left(t_{i|}|t_0, \dots, t_{i-1}, \boldsymbol{\theta}\right) \right] S_{m+1}(a|t_0, \dots, t_n, \boldsymbol{\theta}) dt_m \dots dt_2 dt_1$$

Only those replacements that are related to the failure history, i.e. T^* and *n*, may be represented by the replacement model. Probabilities for independent replacement cancel out in the fraction of (16).

4.2.3 Joint likelihood

If only data of active pipes are available, the likelihood of the pipe failure model and the replacement model must be combined to infer the parameters of the pipe failure model correctly.

The likelihood of the pipe failure model must be conditioned on the event 'pipe has not been replaced up to time *b*'. So the likelihood for an *observed* pipe with *n* recorded failures at times T^* becomes $p(\mathrm{T}^*, n|a, b, 'not\ rep.', \boldsymbol{\theta})$. Expressed according to the Bayes' theorem, this is

$$p(\mathrm{T}^*, n|'not\ rep.', a, b, \boldsymbol{\theta}) = \frac{p(\mathrm{T}^*, n|a, b, \boldsymbol{\theta}) Prob('not\ rep.'|\mathrm{T}^*, n, a, b, \boldsymbol{\theta})}{Prob('not\ rep.'|a, b, \boldsymbol{\theta})} \tag{16}$$

The numerator is the product of the likelihood of the pipe failure model and the replacement model. The denominator of (16) is the probability that a pipe of age *b* has not been replaced, which is obtained by marginalization

$$Prob('not\ rep.'\,|a, b, \boldsymbol{\theta}) = \sum_{n=0}^{\infty} \int_{a}^{b} \int_{t_1^*}^{b} \dots \int_{t_{(n-1)}^*}^{b} p\,(\mathrm{T}^*, n|a, b, \boldsymbol{\theta})$$
$$Prob('not\ rep.'\,|\mathrm{T}^*, n, a, b, \boldsymbol{\theta})\, dt_n^* \dots dt_2^* dt_1^*$$

To obtain the joint likelihood, the likelihoods of the single pipes are multiplied if they are independent.

$$p(\mathrm{T}_1^*, \dots, \mathrm{T}_K^*, n_1, \dots, n_k | 'not\ rep.', a, b, \boldsymbol{\theta}) = \prod_{k=1}^{K} p(\mathrm{T}_K^*, n_k | 'not\ rep.', a, b, \boldsymbol{\theta}) \tag{17}$$

Independence is a reasonable assumption if the pipes are aggregated to a sufficient length (see e.g. Gangl, 2008).

4.2.4 Consideration of covariables

Up to this point, pipes were not distinguished by their properties such as their diameter or material. The same parameter vector $\boldsymbol{\theta}$ was used for all pipes. Consideration of pipe properties can help to improve the predictions for a specific pipe or pipe group and enables the identification of important deterioration processes. Covariables are incorporated by calculating "individual" parameters $\boldsymbol{\theta}_k$ for each pipe *k* as a function of their properties x_k:

$$\boldsymbol{\theta}_k = f(x_k, \boldsymbol{\theta}, \gamma) \tag{18}$$

where γ are additional parameters of $f(\cdot)$ that must be estimated together with $\boldsymbol{\theta}$. To include qualitative pipe properties (e.g. material) indicator variables are used.

4.2.5 Parameter inference

Two widely applied approaches to estimate the parameters are frequentist maximum likelihood estimation (MLE) and Bayesian inference. MLE (e.g. Kleiner and Rajani, 2010; Le Gat, 2009) and Bayesian inference (Dridi et al., 2009; Economou et al., 2009; Watson et al., 2004) have frequently been applied for pipe failure models.

The ML estimator is the parameter vector $\widehat{\boldsymbol{\theta}}$ that maximizes the likelihood function.

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} p(\mathrm{T}_1^*, \dots, \mathrm{T}_K^*, n_1, \dots, n_k | 'not\ rep.', a, b, \boldsymbol{\theta})$$

Large sample properties allow an approximation of the parameter uncertainty (Harrell, 2001).

With Bayesian inference, the distribution of the parameters is calculated given the data and the prior distribution of the parameters $p(\boldsymbol{\theta})$. The prior distribution reflects knowledge about the parameters before the calibration. The proportional relationship (Bernardo and Smith, 2000) is sufficient for numerical calculations:

$$p(\boldsymbol{\theta} | \mathrm{T}_1^*, \dots, \mathrm{T}_K^*, n_1, \dots, n_k, 'not\ rep.', a, b) \propto p(\mathrm{T}_1^*, \dots, \mathrm{T}_K^*, n_1, \dots, n_k | 'not\ rep.', a, b, \boldsymbol{\theta})\, p(\boldsymbol{\theta}) \tag{19}$$

4.2.6 Predictions

In the following the predictive distribution of the number of failures is derived for new pipes and for pipes with a known failure record. Future replacement is purposely not considered in the predictions, to enable the comparison of replacement strategies. The 'pure' predicted failures can then be used directly as input for different replacement strategies.

For the sake of more compact notation, the following predictive distributions are conditioned on the parameters $\boldsymbol{\theta}$. Typically, they will be multiplied by the posterior parameter distribution (19) and then marginalized over $\boldsymbol{\theta}$.

Unconditional predictions

The predictive distribution for a pipe without a failure record is given by the likelihood (13). Typically, interest is limited to the distribution of the number of failures until age *c* which is obtained by marginalization of likelihood (13).

$$Prob(n|c,\boldsymbol{\theta}) = \int_{t_0}^{c}\int_{t_1}^{c}\dots\int_{t_{(n-1)}}^{c} p\,(\mathrm{T},n|c,\boldsymbol{\theta})dt_n dt_{n-1}\dots dt_1 \tag{20}$$

Conditional predictions

To predict the future failures of an existing pipe the failures during the observation period must be considered. Therefore we distinguish between the $n^{(1)}$ observed failures at $T^{*(1)}$ and the $n^{(2)}$ future failures at $\mathrm{T}^{*(2)}$. The predictive distribution of $T^{*(2)}$ and $n^{(2)}$ can be expressed by the likelihood for partially observed failures (14)

$$p\big(\,\mathrm{T}^{*(2)},n^{(2)}|\mathrm{T}^{*(1)},n^{(1)},a,b,c,\boldsymbol{\theta}\big) = \frac{p\big(\,\mathrm{T}^{*(1)}\cup\mathrm{T}^{*(2)},n^{(1)}+n^{(2)}|a,b=c,\boldsymbol{\theta}\big)}{p(\,\mathrm{T}^{*(1)},n^{(1)}|a,b,\boldsymbol{\theta})} \tag{21}$$

The condition '*notrep.*' is not required as it cancels out algebraically. Finally, the distribution of the number of future failures is given by

$$\begin{aligned} &Prob\big(\,n^{(2)}|\mathrm{T}^{*(1)},n^{(1)},a,b,c,\boldsymbol{\theta}\big) \\ &\quad= \int_{b}^{c}\int_{t_1^{(2)}}^{c}\dots\int_{t_{(n-1)}^{(2)}}^{c} p\big(\,\mathrm{T}^{*(2)},n^{(2)}|\mathrm{T}^{*(1)},n^{(1)},a,b,c,\boldsymbol{\theta}\big)\,dt_{n^{(2)}}^{(2)}dt_{(n^{(2)}-1)}^{(2)}\dots dt_1^{(2)} \end{aligned} \tag{22}$$

4.3 Example: Weibull-exponential model

While the general description above provides the 'recipe' for the likelihood for a particular model, there is no assurance that the resulting likelihood can be handled algebraically and numerically. In this section, we show how the likelihood for a rather simple pipe failure model that was applied by Mailhot et al. (2000) can be combined with an elementary replacement model.

For the pipe failure model, we assume that the time from construction until the first failure is Weibull distributed, and the time between all following failures exponential with the same rate parameter.

This failure model requires three parameters: the shape parameter θ_1 and the scale θ_2 of the Weibull distribution

$$\begin{aligned} p_1(t|t_0,\boldsymbol{\theta}) &= \frac{\theta_1}{\theta_2}\left(\frac{t-t_0}{\theta_2}\right)^{\theta_1-1} e^{-[(t-t_0)/\theta_2\,]^{\theta_1}} \\ S_1(t|t_0,\boldsymbol{\theta}) &= e^{-[(t-t_0)/\theta_2\,]^{\theta_1}} \end{aligned} \tag{23}$$

and the scale θ_3 of the exponential distribution

$$p_i(t|t_0,\dots,t_{i-1},\boldsymbol{\theta}) = p_i(t|t_{i-1},\boldsymbol{\theta}) = \frac{1}{\theta_3}e^{-(t-t_{i-1})/\theta_3}$$
$$S_i(t|t_0,\dots,t_{i-1},\boldsymbol{\theta}) = S_i(t|t_{i-1},\boldsymbol{\theta}) = e^{-(t-t_{i-1})/\theta_3} \tag{24}$$

for all $i > 1$.

4.3.1 Likelihood for completely observed pipes

The distributions defined in (23) and (24) are directly plugged into the general likelihood for completely observed pipes (13). After some algebraic rearrangements, we obtain

$$p(\mathrm{T},n|b,\boldsymbol{\theta}) = \begin{cases} e^{-[(t-t_0)/\theta_2\,]^{\theta_1}}, & n=0 \\ \dfrac{\theta_1}{\theta_2}\left(\dfrac{t_1-t_0}{\theta_2}\right)^{\theta_1-1} e^{-[(t_1-t_0)/\theta_2\,]^{\theta_1}}\left(\dfrac{1}{\theta_3}\right)^{n-1} e^{-(b-t_1)/\theta_3}, & n>0 \end{cases} \tag{25}$$

Note that due to the algebraic form of the exponential distribution and the assumption that the rate parameter remains the same for all $i > 1$ the likelihood only depends on the time of the first failure and on the number of failures *n*.

Likelihood for partly observed pipes

Similarly, the likelihood for partly observed pipes must be distinguished for $n = 0$ and $n > 0$. If no failures are observed ($n = 0$), it is

$$p(\mathrm{T}^*,n=0|a,b,\boldsymbol{\theta}) = e^{-[(b-t_0)/\theta_2\,]^{\theta_1}} + e^{-(b-a)/\theta_3}\left[1-e^{-[(a-t_0)/\theta_2\,]^{\theta_1}}\right] \tag{26}$$

and for *n*>0

$$p(\mathrm{T}^*,n|a,b,\boldsymbol{\theta}) = \left(\frac{1}{\theta_3}\right)^{n-1} e^{-(b-t_1^*)/\theta_3}\left\{\frac{\theta_1}{\theta_2}\left(\frac{t_1^*-t_0}{\theta_2}\right)^{\theta_1-1} e^{-[(t_1^*-t_0)/\theta_2\,]^{\theta_1}} + \frac{1}{\theta_3}e^{(a-t_0)/\theta_3}\left[1-e^{-[(a-t_0)/\theta_2\,]^{\theta_1}}\right]\right\} \tag{27}$$

See Mailhot et al. (2000) or Pelletier (2000) for a derivation of these equations.

4.3.2 Consideration of replacement

The replacement model is first defined conditioned on *m*. The simplest model with a single parameter assumes a constant probability π that a pipe is not replaced if a failure occurs: $Prob('not\ rep.'\,|\mathrm{T}^*,n,m,a,b,\boldsymbol{\theta}) = \pi^{m+n}$. In the following we assume that π is contained in the parameter vector Θ. The resulting unconditional replacement model is then

$$Prob('not\ rep.'\,|\mathrm{T}^*,n,a,b,\boldsymbol{\theta}) = \sum_{m=0}^{\infty}\pi^{m+n}Prob(m|a,\boldsymbol{\theta}) \tag{28}$$

where the probability of *m* unobserved failures before *a* is

$$Prob(m|a,\boldsymbol{\theta}) = \begin{cases} e^{-\left[\frac{(a-t_0)}{\theta_2}\right]^{\theta_1}}, & m=0 \\ \int_{t_0}^{a}\int_{t_1}^{a}\dots\int_{t_{(m-1)}}^{a} \frac{\theta_1}{\theta_2}\left(\frac{t_1-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1-t_0}{\theta_2}\right]^{\theta_1}} \left(\frac{1}{\theta_3}\right)^{m-1} e^{-\frac{(a-t_1)}{\theta_3}} dt_m \dots dt_2 dt_1, & m>0 \end{cases}$$

$$= \begin{cases} e^{-\left[\frac{t-t_0}{\theta_2}\right]^{\theta_1}}, & m=0 \\ \frac{\theta_1}{\theta_2}\int_{t_0}^{a}\left(\frac{t_1-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1-t_0}{\theta_2}\right]^{\theta_1}} \left(\frac{1}{\theta_3}\right)^{m-1} e^{-\frac{(a-t_1)}{\theta_3}} \frac{(a-t_1)^{m-1}}{(m-1)!} dt_1, & m>0 \end{cases}$$

The integrand depends only on t_1 so the remaining integrals reduce to $\int_{t_1}^{a}\dots\int_{t_{(m-1)}}^{a} 1\, dt_m \dots dt_2 = \frac{(a-t_1)^{m-1}}{(m-1)!}$.

Combined with the conditional replacement model, the sum and the factorials can be simplified by recognizing that $\sum_{k=0}^{\infty}\frac{x^k}{k!} = e^x$ (see also Pelletier et al.(2003), page 83):

$$\begin{aligned} &Prob('not\ rep.'\,|\mathrm{T}^*, n, a, b, \boldsymbol{\theta}) \\ &= e^{-\left[\frac{(a-t_0)}{\theta_2}\right]^{\theta_1}} \\ &+ \pi^{n+1}\frac{\theta_1}{\theta_2}\int_{t_0}^{a}\left(\frac{t_1-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1-t_0}{\theta_2}\right]^{\theta_1}} e^{-\frac{(b-t_1)}{\theta_3}} \sum_{m=1}^{\infty}\left(\frac{1}{\theta_3}\right)^{m-1} \pi^{m-1}\frac{(a-t_1)^{m-1}}{(m-1)!} dt_1 \\ &= e^{-\left[\frac{(a-t_0)}{\theta_2}\right]^{\theta_1}} + \pi^{n+1}\frac{\theta_1}{\theta_2}\int_{t_0}^{a}\left(\frac{t_1-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1-t_0}{\theta_2}\right]^{\theta_1}} e^{-\frac{(b-t_1)}{\theta_3}} e^{\pi\frac{(a-t_1)}{\theta_3}} dt_1 \end{aligned} \tag{29}$$

The replacement model (29), and the likelihood of the pipe failure model for partly observed pipes without (26) or with failures (27) are combined in (16) to obtain the conditional likelihood. Although the integrals of the denominator cannot be solved analytically, it can be simplified sufficiently (details not shown) to allow numerical integration without problems:

$$\begin{aligned} &Prob('not\ rep.'\,|a, b, \boldsymbol{\theta}) \\ &= e^{-\left[\frac{(b-t_0)}{\theta_2}\right]^{\theta_1}} + \frac{\theta_1}{\theta_2}\int_{t_0}^{a}\left(\frac{t_1-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1-t_0}{\theta_2}\right]^{\theta_1}} e^{-\frac{(b-t_1)}{\theta_3}} \pi e^{\pi\frac{(a-t_1)}{\theta_3}} dt_1 \\ &+ \frac{\theta_1}{\theta_2}\int_{a}^{b}\left(\frac{t_1^*-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1^*-t_0}{\theta_2}\right]^{\theta_1}} e^{-\frac{(b-t_1^*)}{\theta_3}} \pi e^{\pi\frac{(b-t_1^*)}{\theta_3}} dt_1^* \\ &+ \frac{\theta_1}{\theta_2}\frac{\pi^2}{\theta_3}\int_{t_0}^{a}\int_{a}^{b} e^{-\frac{(b-t_1^*)}{\theta_3}} e^{\pi\frac{(b-t_1^*)}{\theta_3}} \left(\frac{t_1^*-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left[\frac{t_1-t_0}{\theta_2}\right]^{\theta_1}} e^{-\frac{t_1^*-t_0}{\theta_2}} e^{\pi\frac{(a-t_1)}{\theta_3}} dt_1^*\, dt_1 \end{aligned} \tag{30}$$

4.3.3 Predictions

Unconditional predictions

Monte Carlo samples can be conveniently generated from the likelihood for completely observed pipes (25). From there, it is straightforward to obtain the distribution of the number of failures by sequentially sampling from the Weibull (23) and exponential distributions (24).

Conditional predictions

It is not trivial, however, to sample from the likelihood for partly observed pipes (21). Instead, an expression proportional to $Prob\left(n^{(2)} \middle| \mathrm{T}^{*(1)}, n^{(1)}, a, b, c, \boldsymbol{\theta}\right)$ can be derived, on whose basis a sample can be obtained with importance or Metropolis sampling. The formulations must be distinguished depending on the number of observed $n^{(1)}$ and predicted $n^{(2)}$ failures.

If $n^{(1)} > 0$:

$$Prob\left(n^{(2)} \middle| \mathrm{T}^{*(1)}, n^{(1)}, a, b, c, \boldsymbol{\theta}\right) \propto p\left(\mathrm{T}^{*(1)}, n \middle| a, b, \boldsymbol{\theta}\right)\Big|_{n=n^{(1)}+n^{(2)}, b=c} \frac{(c-b)^{n^{(2)}}}{n^{(2)}!}$$

and if no failures were observed, i.e. $n^{(1)} = 0$:

$$Prob\left(n^{(2)} \middle| \mathrm{T}^{*(1)}, n^{(1)}, a, b, c, \boldsymbol{\theta}\right) \propto \begin{cases} p(\mathrm{T}^*, n=0 | a, b, \boldsymbol{\theta})|_{b=c} & n^{(2)} = 0 \\ \int_b^c p(\mathrm{T}^*, n | a, b, \boldsymbol{\theta})|_{n=n^{(2)}, b=c} \frac{(c-t_1^*)^{n^{(2)}-1}}{(n^{(2)}-1)!} dt_1^* & n^{(2)} > 0 \end{cases}$$

For $p(\mathrm{T}^*, n | a, b, \boldsymbol{\theta})$and $p(\mathrm{T}^*, n=0 | a, b, \boldsymbol{\theta})$ see (26) and (27) for partly observed pipes, respectively.

4.3.4 Implementation

Procedures for inference and prediction were implemented in R (R Development Core Team, 2011) that evokes a Fortran 95 implementation of the likelihood function. Samples of the posterior were obtained with the adaptive Metropolis sampler proposed by Vihola (2011) and implemented in the R-package *adaptMCMC* (Scheidegger, 2012).

4.4 Application examples

4.4.1 Simulated data

Two data sets are simulated to show that replacement causes biased parameter estimations if not considered appropriately.

The data sets have different sample sizes: they consist of the failure records of 100 and 1 000 pipes respectively. The failures were generated on the basis of the distribution assumptions made for the failure model in section 4.3 (time to first break is Weibull distributed, time between the following breaks exponential). Replacement was simulated according to the replacement model of section 4.3.2. The data sets were then compiled from failures within the observation period of the unreplaced pipes only. Data sets for a 60 years old system were simulated (for parameter values, see Table 16). The recording period was assumed to cover the last 10 years; 39 failures occurred in this period for the small data set and 403 for the large one.

The parameters are inferred with two models: a) the failure model of section 4.3 while ignoring replacement, and b) the same failure model combined with the replacement model of section 4.3.2. The prior distribution and the summarized posterior based on seven Monte Carlo Markov chains with 100 000 samples each are shown in Table 16 for both models.

The expected number of failures, calculated with equation (20) for unconditional predictions, as a function of the pipe age for newly built pipes, is shown in Figure 9. On Figure 9a) it is clearly seen that the failure frequency is underestimated if the replacement is not considered for the parameter estimation. With increasing sample size only the uncertainty becomes smaller, while the bias remains constant. Figure 9b) shows the result for parameters estimated with the replacement model. Although the uncertainties are greater than in Figure 9a), no systematic deviation is present.

Table 16: Parameters used for data simulation and estimated values. The posterior is based on seven independent MCMC chains with 100 000 samples each. The marginal posterior distributions are summarized with three numbers: first the mean followed by the 10 % and 90 % quantiles.

	Data generation	Prior U(l, u)	Posterior without replacement	Posterior with replacement
100 pipes				
θ_1	2	(0.5, 4)	3.03 (2.11, 3.82)	3.05 (2.16, 3.82)
θ_2	30	(1, 250)	36.75 (29.97, 44.90)	31.99 (24.79, 39.60)
θ_3	15	(1, 250)	20.81 (15.43, 26.90)	15.20 (8.36, 22.18)
π	0.75	(0, 1)	–	0.73 (0.44, 0.96)
1 000 pipes				
θ_1	2	(0.5, 4)	2.28(2.03, 2.54)	2.23 (1.97, 2.49)
θ_2	30	(1, 250)	33.58(31.14, 36.13)	30.26 (25.76, 34.14)
θ_3	15	(1, 250)	19.72(18.14, 21.37)	15.84 (11.29, 19.52)
π	0.75	(0, 1)	–	0.81 (0.59, 0.97)

Figure 10 shows all one- and two-dimensional marginals of the posterior parameter distribution. The one-dimensional marginals reveal the shape of the posterior for each parameter. Dependencies are visible in the two-dimensional marginals. So is a clear correlation apparent between the replacement probability and the two scale parameters. This is important for real applications as it implies that an informative prior for the replacement probability would reduce the uncertainty of the scale parameters considerably.

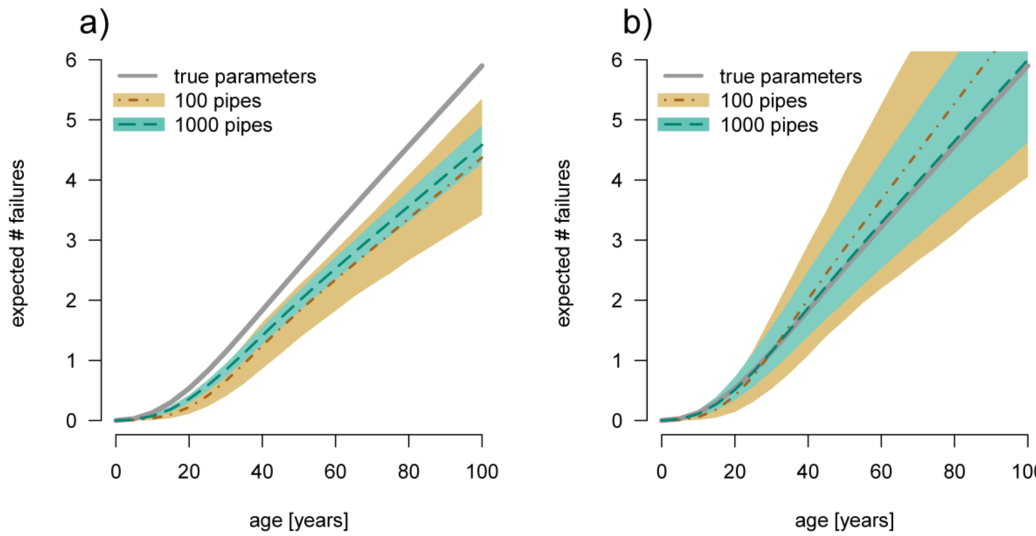


Figure 9: Estimated expected number of failures for a pipe as function of age. The estimations are based on simulated data with 100 and 1 000 pipes respectively. For Figure a) the parameters were estimated without consideration of a replacement model. Figure b) shows the estimation with the replacement model. The dashed lines show the mean and the shaded areas indicate the 80 %-credibility interval of the expected number of failures.

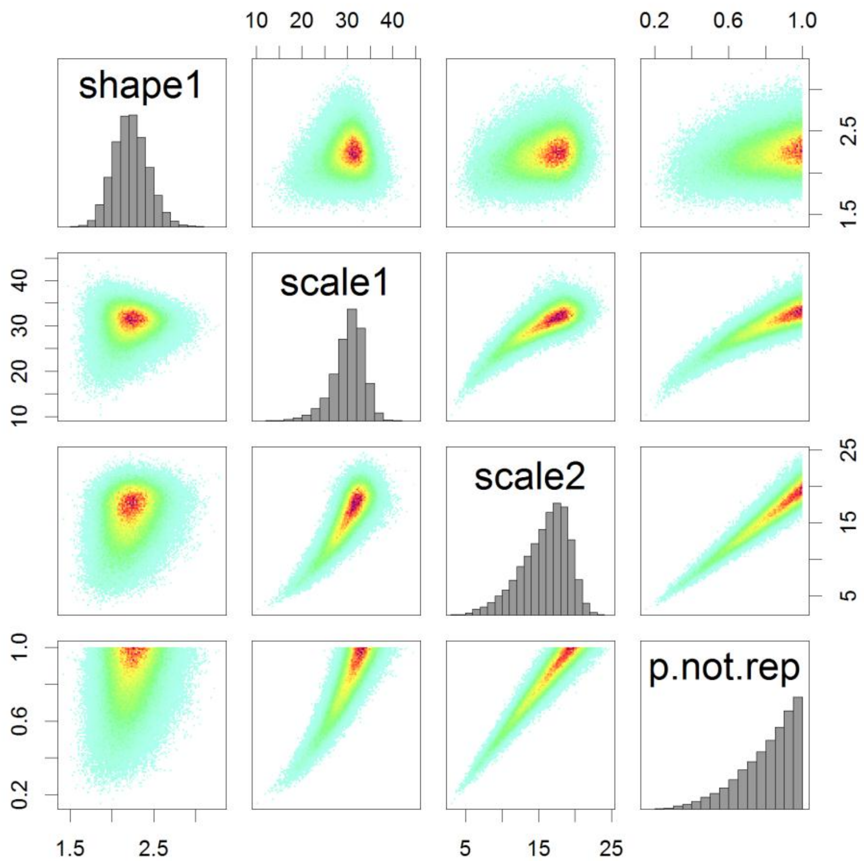


Figure 10: Marginals of the posterior distribution based on simulated data of 1 000 pipes and uniform priors. Warm colors denote regions with high probability density.

4.4.2 Real data

The second example illustrates one way of modifying the model to incorporate covariables. Covariables can represent quantitative (diameter, …) or qualitative (material, construction period, …) properties of an individual pipe. In this example, the influence of the construction period of ductile cast iron pipes is investigated.

Data on the water supply of Lausanne, Switzerland, is used. Instead of using the whole network, the focus is on one characteristic material only, namely, ductile cast iron (DI). It makes up the largest proportion (about 62 %) of the network and can be divided into two generations according to the manufacturing and laying periods: pipes with rather poor protection against outer corrosion (DI1), and pipes with improved corrosion protection (DI2). In Switzerland, DI1 pipes were commonly used until 1980 and were then succeeded by DI2. The proportion of DI2 pipes with recorded failures until the present is low (about 2 %), often making parameter inference a challenge. To reduce the influence of pipe length on the modeling of first and subsequent failures (Fuchs-Hanusch et al., 2012; Gangl, 2008; Poulton et al., 2007), the approach of Gangl (2008) was used. Gangl (2008) suggests forming 100 to 200 m long pipe units from neighboring pipes with equal diameters, materials and laying years. This is based on an analysis of the distances between subsequent failures, which are usually below 100 m and no longer than 200 m. Thus, short pipe segments were merged to 444 segments of DI1 pipes and to 2 636 segments of DI2 pipes. As no spatial information was available, the merging was based on the construction year and diameter only. The average length of the merged segments is 143.1 m. Pipe failures were systematically recorded over ten years. Failures within the first year after installation were removed, because they are attributed to installation deficiencies and not to structural aging. The record contains then 116 failures of DI1 pipes and 82 failures of DI2 pipes.

The qualitative information about the construction period is modeled with the help of indicator variables. For each pipe *k*, "individual" parameters $\boldsymbol{\theta}_k$ are computed as described in equation (18). In this case we choose $f(\cdot)$as

$$\boldsymbol{\theta}_k = f(x_k, \boldsymbol{\theta}, \boldsymbol{\gamma}) = (\theta_1, \gamma^{x_k}\theta_2, \gamma^{x_k}\theta_3, \theta_4)^{\boldsymbol{T}}$$

where x_k is the indicator variable that equals one if pipe *k* is a DI2 pipe and zero otherwise. Accordingly, θ_2 and θ_3 can be interpreted as scale parameters for DI1 pipes and $\gamma\theta_2$ and $\gamma\theta_3$ as scales for DI2 pipes.

The same uniform priors as for the first example (Table 16) were used for $\boldsymbol{\theta}$, and a gamma distribution for $\boldsymbol{\gamma}$ with mode one and a standard deviation of five.

As in the previous example we inferred the parameters without (Figure 11a) and with (Figure 11b) consideration of the replacement model. The resulting expected number of failures as function of pipe age is shown for both pipe generations in Figure 11a) and 11b). The corresponding posterior parameter distributions are summarized in Table 17.

As expected, the model predicts a higher failure rate if it corrects for pipe replacement. However, the predictions have larger uncertainties due to the additional parameter. The probability π that a pipe is not replaced after a failure is estimated within a reasonable range (see Table 17). Pipes of the first generation have a considerably higher risk of failures. This is in line with observations from practice in Switzerland; the lack of corrosion protection and the grounding of electrical appliances on the water lines until the 1990s led to increased corrosion and a large number of failures of DI1 pipes (Kappeler et al., 2010). Differences in the failure behavior of pipes of different installation periods were frequently observed (e.g.

Kleiner and Rajani, 1999; Mailhot et al., 2000). Because the data contain fewer DI1 pipes, the estimation is more uncertain than that of DI2.

Table 17: The resulting posterior parameter distributions for data from Lausanne inferred with and without replacement model. The posteriors are based on seven independent MCMC chains with 100'000 samples each. The posterior marginal distributions are summarized by three numbers: first the mean followed by the 10 % and 90 % quantiles.

	without repl. model	with repl. model
$\boldsymbol{\theta}_1$	1.80 (1.52, 2.10)	1.80 (1.53, 2.09)
$\boldsymbol{\theta}_2$	73.47 (63.82, 84.56)	65.88 (54.88, 77.35)
$\boldsymbol{\theta}_3$	14.55 (11.98, 17.39)	13.05 (10.20, 16.08)
γ	1.80 (1.50, 2.13)	1.84 (1.52, 2.20)
π	–	0.89 (0.76, 0.99)

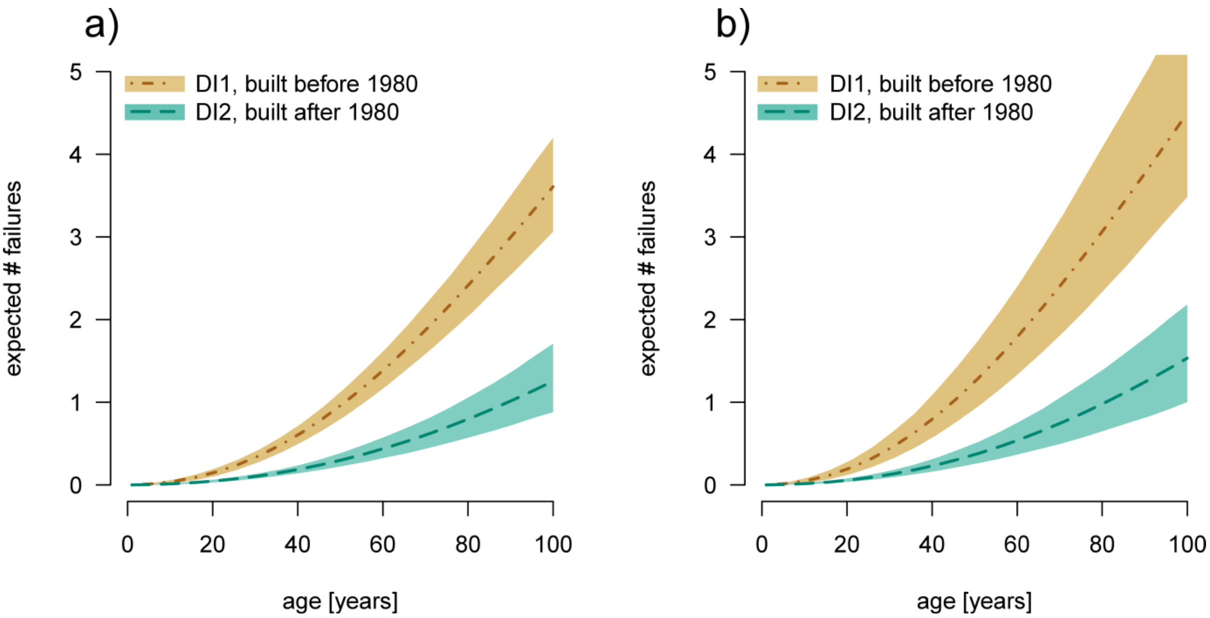


Figure 11: Expected number of failures for two generations of ductile iron pipes of the Lausanne water supply network. For Figure a) the parameters were estimated without consideration of a replacement model. Figure b) shows the estimation with the replacement model. The dashed lines show the mean and the shaded areas indicate the 80%-credibility interval of the expected number of failures.

4.5 Discussion

We demonstrated in Section 4.2 how a failure based pipe model can be extended by a replacement model. It is important to realize, that we did not propose a new pipe failure model. The intention was to present a procedure to avoid selective survival biases by adopting existing (or future) models. Therefore, we first introduced a generic notation, exceeding the "general framework for water main break modeling" of Mailhot et al. (2003). Furthermore, we also derived all the equations for left truncated observations, including the likelihood function and the predictive distribution.

Our approach is limited to pipe models that (a) are failure based, (b) consider—if known—the failure history of a specific pipe for predictions, and (c) allow probabilistic statements about parameter and prediction uncertainty, i.e models based on a likelihood function. Condition (a) excludes life-span models (Herz, 1995; Scholten et al., 2013) as they intentionally lump failure behavior and the often consequential replacement together. Regression models (Boxall et al., 2007; Kleiner and Rajani, 1999) and simple proportional hazard (Cox) models (Carrión et al., 2010; Gangl, 2008) do not fulfill condition (b) and are therefore of limited use for prediction. Purely data driven algorithms (Giustolisi et al., 2006; Jafar et al., 2010) often do not fulfill condition (c) and therefore do not fit into a probabilistic framework. Models excluded here may nevertheless be influenced by the survival selection bias.

In Section 4.3 we exemplified our approach with a Weibull-exponential model. This model was chosen because it has been successfully applied (Mailhot et al., 2000) and has manageable complexity. However, as any parametric distribution the Weibull has some limitations. In particular, the hazard rate begins at zero (if shape parameter >1) and therefore installation failures and the probability of third party damages (typically caused by construction activities, Thomson and Wang, 2009) of young pipes cannot be modeled.

We demonstrated in the first example with artificially generated data that for datasets without historic data (containing information about replaced pipes) consideration of a replacement model is crucial for reliable predictions. In these cases, ignoring the replacement of pipes leads to a bias that cannot be reduced by increasing the sample size. This is especially significant in well maintained networks in which substantial replacements were made in the past. In these cases the failure rates are strongly underestimated. Missing data on replaced pipes is very common for many networks that we encounter here in Switzerland, and we suspect that it is equally common elsewhere (e.g. Le Gat, 2009).

The second example was based on a data set of ductile iron pipes from a real water supply network. The results show the expected behavior: (i) the first generation ductile pipes have a clearly higher failure rate, and (ii) the dataset with fewer observations shows larger uncertainties. This also illustrates a possible approach to extending the model by covariables. They give the model more flexibility to fit the data. An alternative is to group pipes in homogeneous sets and fit a model independently to each of those. However, incorporating covariables has the advantage that interactions can be revealed and that, in total, fewer parameters need be inferred.

Inevitably, the likelihood function becomes more complicated if a replacement model is included, in particular if the data are left truncated (common for many European water networks). However, a replacement model may not be required, if data of replaced pipes are available. For some models the representation as counting process is more practicable. Instead of translating such models into the time domain, it might be more feasible to modify the presented approach accordingly. It is important to realize that the data availability and ultimately the data collection scheme determines the correct likelihood function. Therefore, the modeler must have a clear understanding of how the data have been collected and managed. This information is usually not directly evident from the data.

Replacement models do not have to represent decisions that are independent of the failure record, for example replacement decisions that are based purely on the age of the pipes. They do not lead to a bias in the parameter inference as terms representing independent decisions cancel out algebraically. Considering such replacement strategies in the replacement model would add unnecessary complexity.

The simplest possible replacement model was chosen for the examples, and it is certainly not suitable for all data sets. Therefore, the development of a more general replacement model could be helpful. For example replacement decisions may depend on the pipe age when failures occur or replacement strategies could change over time. Furthermore, in the context of Bayesian inference, how the prior distribution of the replacement model parameters is elicited optimally should be investigated. A good elicitation method to obtain informative priors from other utilities and / or expert elicitation would be critical for applications to small utilities with scarce data.

Selecting the most appropriate model remains a challenging task that cannot be automated. In many utilities, the data handling was guided by daily operation requirements and not with failure modeling in mind. This results in data sets that require the application of an adequate model, suited to the particular data management characteristics. This problem is not exclusive to drinking water pipes. For example, a similar approach to sewer deterioration models is developed by (Egger et al., 2013). Adapting the failure models to specific data sets is the only option, in the absence of more standardized data management strategies.

4.6 Conclusions

- Pipe failure models are important tools for the management of water distribution networks. The calibration of such models is often complicated by common practices of data handling. A frequent problem is that many available data sets contain records of pipes in service but not of replaced pipes. Calibration without explicit accounting for this practice can lead to considerably biased predictions.

- To correct for such biases, we propose an approach to modify the likelihood function of failure models. The key idea is to combine the likelihood function of the failure model with a probabilistic replacement model.

- As past replacement and data management practices are different for every network, a failure model must be adapted to a specific data set. The approach presented here is formulated generally and is therefore applicable to many—existing or future—pipe failure models and different replacement models.

- The concept is illustrated explicitly for a Weibull-exponential model in combination with a simple replacement model. Furthermore, we show how models can be extended to consider covariables.

- The code of the model used in the examples is freely available on request.

4.7 Acknowledgments

We thank Sébastien Apothéloz and Eauservice Lausanne for their cooperation and contribution of data.

5 Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis

Lisa Scholten[a,b)], Andreas Scheidegger[a)], Peter Reichert[a,b)], Max Mauer[a,c)], Judit Lienert[a)]

[a)] Eawag: Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, P.O.Box 611,CH-8600 Duebendorf, Switzerland
[b)] ETH Zurich, Institute of Biochemistry and Pollutant Dynamics (IBP), CH-8092 Zurich, Switzerland
[c)] ETH Zurich, Institute of Environmental Engineering, CH-8093 Zurich, Switzerland

Water Research (2014) , 49, 124-143. http://dx.doi.org/10.1016/j.watres.2013.11.017

Abstract
To overcome the difficulties of strategic asset management of water distribution networks, a pipe failure and a rehabilitation model are combined to predict the long-term performance of rehabilitation strategies. Bayesian parameter estimation is performed to calibrate the failure and replacement model based on a prior distribution inferred from three large water utilities in Switzerland. Multi-criteria decision analysis (MCDA) and scenario planning build the framework for evaluating 18 strategic rehabilitation alternatives under future uncertainty. Outcomes for three fundamental objectives (low costs, high reliability, and high intergenerational equity) are assessed. Exploitation of stochastic dominance concepts helps to identify twelve non-dominated alternatives and local sensitivity analysis of stakeholder preferences is used to rank them under four scenarios. Strategies with annual replacement of 1.5-2 % of the network perform reasonably well under all scenarios. In contrast, the commonly used reactive replacement is not recommendable unless cost is the only relevant objective. Exemplified for a small Swiss water utility, this approach can readily be adapted to support strategic asset management for any utility size and based on objectives and preferences that matter to the respective decision makers.

Keywords
Strategic water asset management, failure and rehabilitation modeling, water supply, multi-criteria decision analysis, decision support, scenario planning

5.1 Introduction

5.1.1 Strategic Asset Management (SAM)

Awareness about the need for long-term rehabilitation planning of our aging water infrastructure has risen globally during the past two decades (AWWA, 2001; Burns et al., 1999; Herz, 1998; Kleiner and Rajani, 1999; Sægrov, 2005; Selvakumar and Tafuri, 2012; Vanier, 2001). Infrastructure asset management (IAM) is increasingly applied to rehabilitation planning on the strategic, tactical, and operational levels (Cardoso et al., 2012; Christodoulou et al., 2008; Fuchs-Hanusch et al., 2008; Haffejee and Brent, 2008; Heather and Bridgeman, 2007; Marlow et al., 2010; Ugarelli et al., 2010).

Recently, the CARE-W (Sægrov, 2005) and AWARE-P (Cardoso et al., 2012) research projects have greatly contributed to the development and implementation of structured IAM approaches, including strategic asset management (SAM). Both rely on (i) knowledge about the expected useable lifetime and condition of assets over time (failure models), (ii) knowledge about the consequences of rehabilitation alternatives (rehabilitation models), but are weak in (iii) systematic and transparent decision support, and (iv) thorough accounting for planning uncertainty.

Application of the available SAM approaches in the water sector is still limited, given the high need for human, informational, and data resources (Alegre, 2010). In Switzerland, SAM is a specific challenge due to the sector's high fragmentation (Lienert et al., 2013) and prevalence of mostly small water providers, the majority with < 10'000 beneficiaries (SVGW, 2006).

5.1.2 Failure models

To compare water network rehabilitation options, knowledge about the expected useable lifetime and condition of pipe assets is crucial (Selvakumar and Tafuri, 2012). Probabilistic water pipe failure models to predict age-dependent pipe deterioration abound (Kleiner et al., 2009; Kleiner and Rajani, 2001; Liu et al., 2012). Whereas their practical value has been shown especially in connection to larger water networks (e.g. Alvisi and Franchini, 2010; Eisenbeis et al., 1999; Poulton et al., 2007; Renaud et al., 2012), their calibration to the local conditions is usually infeasible in small to medium-sized water networks because of their high data demand. Hence, there is a lack of failure models that support rehabilitation planning in the very common small to medium-sized networks in Switzerland, but also in other European countries such as Austria, Germany, and France. Additionally, common data particularities, namely left-truncation, right-censoring, and selective survival bias, are usually not explicitly considered in model parameter inference, which may lead to biased predictions of failures (Le Gat, 2009; Mailhot et al., 2000; Renaud et al., 2012; Scheidegger et al., 2011). A general approach as well as a specific model to avoid biases in pipe failure models due to these particularities were recently proposed by Scheidegger et al. (2013). The problem of short networks (small sample size) and limited failure records in pipe failure model calibration can be overcome by Bayesian parameter inference (Dridi et al., 2009; Watson et al., 2004).

5.1.3 Comparing rehabilitation alternatives

The available rehabilitation models are mostly used to support operational and tactical (i.e. short to mid-term) pipe repair and replacement planning (for a review see Engelhardt et al., 2000). Nonetheless, software to support strategic (long-term) rehabilitation decisions exists, usually combining pipe deterioration and evaluation models with decision support features (e.g. KANEW (Kropp and Baur, 2005), PiReM (Fuchs-Hanusch et al., 2008), D-WARP (Kleiner and Rajani, 2004), Aware-P (Cardoso et al., 2012), Casses (Renaud et al., 2012), WilCO (Engelhardt et al., 2003), PARMS Planning (Burn et al., 2003)). From the information available, and examining four software products in detail, we judged none suitable to simultaneously meet core requirements of our approach: a) combinability with our failure model, b) flexible implementation of rehabilitation strategies and performance measures, and c) propagation of parameter uncertainty. We therefore selected the sector-independent asset management software FAST (Fichtner Asset Services & Technologies, 2013) which is based on a set of interacting differential equations as used in system dynamic modeling. E.g. Rehan et al. (2011) follow a system dynamic approach for the long-term planning of water and wastewater systems and studying the financial sustainability of different rehabilitation strategies.

5.1.4 Decision support

As noted by others, e.g. (Alegre, 2010; Giustolisi et al., 2006; Selvakumar and Tafuri, 2012), the evaluation and prioritization of water system rehabilitation alternatives should be supported by robust and feasible decision support tools. In water engineering, single- or multi-objective optimization and cost-benefit analysis are commonly used to support decisions (Engelhardt et al., 2000; Giustolisi et al., 2006) although they often ignore subjective stakeholder preferences. In a long-term and multi-stakeholder context like strategic rehabilitation planning, the integration of stakeholder preferences by multi-criteria decision analysis (MCDA) seems more appropriate (Keeney, 1982).

MCDA has been applied to water infrastructure asset management at least twice (Baur et al., 2003; Carriço et al., 2012); both using ELECTRE of the outranking family of MCDA methods (Roy, 1991). Many other MCDA approaches are available, see e.g. Belton and Stewart (2002) and Figueira et al. (2005) for an overview. Another well-established MCDA approach is multi-attribute value and utility theory (MAVT / MAUT). Four important reasons for choosing MAVT / MAUT to support asset management decisions (further explained in Schuwirth et al., 2012a) are: 1) foundation on axioms of rational choice, 2) explicit handling of prediction uncertainty and stakeholder risk attitudes, 3) ability to process many alternatives without increased elicitation effort, and 4) possibility to include new alternatives at any stage of the decision procedure.

5.1.5 Uncertainty assessment

A major concern for long-term planning is the consideration of uncertainty about future developments, the probabilistic description of which is difficult due to high ambiguity (Rinderknecht et al., 2012). Scenario planning has been proposed to handle these uncertainties (Schnaars, 1987) and mitigate under- and over- prediction of change (Schoemaker, 1995). It is increasingly incorporated into both IAM and MCDA to evaluate the robustness of decision alternatives to future change (Cardoso et al., 2012; Goodwin and Wright, 2001; Karvetski et al., 2009b; Montibeller et al., 2006; Stewart et al., 2013). While scenario thinking can be interpreted as a way to cover in-between uncertainties of a range of possible futures, uncertainty quantification and propagation of model outputs combined with sensitivity analysis allows the consideration of uncertainty within future scenarios (Stewart et al., 2013).

5.1.6 Goal and structure

Recent reports confirm that the need for water infrastructure rehabilitation in Switzerland is higher than actual rehabilitation (Martin, 2009), but strategic planning is missing. Higher rehabilitation needs have also been recognized in other places, e.g. Australia (Burns et al., 1999), and the USA (Selvakumar and Tafuri, 2012). Our main objective is to show ways out of this planning backlog. We demonstrate a novel approach on how long-term rehabilitation strategies can be evaluated by integrating failure and rehabilitation modeling into a multi-criteria decision analysis (MCDA) and scenario planning framework. We aim at answering two key questions:

1. Which outcomes are expected for different pipe rehabilitation strategies?
2. Which are the best rehabilitation strategies under given preferences and how robust are they under different future scenarios?

A small Swiss water utility ("D") serves as practical example to illustrate that SAM is possible even in small utilities. The deterioration model and its calibration are geared to small networks and can be replaced by other approaches depending on the amount of data available and the desired sophistication of failure modeling. The overall MCDA approach, however, should scale well for any utility size.

The remainder of this manuscript is organized as follows: In section 5.2.1, a new length homogenization procedure is presented to allow the comparison of four water networks, A-D. Secondly, parameters for the failure model are estimated for networks A-C and aggregated into one prior parameter distribution (5.2.2). The posterior failure parameters for D are obtained by Bayesian inference; failures before the start of failure recording in D are also predicted. Thirdly, the posterior parameters from (5.2.2) are inputs to model the outcomes of 18 rehabilitation alternatives under four future scenarios by means of a rehabilitation model (5.2.3) for utility D. Fourthly, the rehabilitation alternatives' outcomes are evaluated with MCDA, assuming different stakeholder preferences (5.2.4-5.2.9). To remove irrelevant alternatives, dominance concepts are exploited. A local sensitivity analysis determines the robustness of the alternatives' ranking to preference changes under future scenarios. Additional information and figures, including a list of symbols and abbreviations, is given in the supporting information (SI)

5.2 Material and methods

5.2.1 Data preparation

Four Swiss water suppliers of different size provided their data to this study. The three larger ones (A-C) are used to infer the Bayesian prior and the smallest is the target utility (D). To facilitate comparison, the pipe and failure data of A-D are prepared in the same manner.

Failures occurring in the installation year are discarded as they are likely caused by installation deficiencies and not structural aging. After plausibility checks, pipes are grouped by shared properties, known to affect pipe deterioration, especially material, date of laying, and diameter (Carrión et al., 2010; Giustolisi et al., 2006; Kleiner and Rajani, 1999). Relevant groups for D are, differentiated by material and laying period: 1st and 2nd generation ductile cast iron (DI1 before, DI2 after 1980; both centrifugal casting, but DI1 only with lacking outer corrosion protection), 2nd and 3rd generation grey cast iron (GI2 before, GI3 after 1930; vertical and centrifugal casting, respectively), asbestos cement incl. Eternit (FC), steel (ST), and polyethylene (PE). In utility D, pipe laying dates of ca. 98 % of pipes were known precisely. For the remaining 2 %, the midpoint of the stated time interval was used. The results from Bayesian inference did not significantly differ when taking the minimum or maximum point of the intervals (not shown), such that uncertainty arising from this was neglected. Further specification of sub-groups into diameter classes or external influences (e.g. road traffic, soil conditions) is avoided in order not to excessively stratify the already few failure data available.

The influence of pipe length on failure prediction is important in failure modeling (Carrión et al., 2010; Fuchs-Hanusch et al., 2012; Gangl, 2008; Poulton et al., 2007), because failures are often triggered by previous failures in the vicinity (Rajani and Kleiner, 2001). One solution would be its explicit consideration as additional model covariate, requiring more parameters to be estimated. Instead, we homogenize the data by merging and splitting, based on the observation of a large Austrian water network (Graz), where roughly 95 % of subsequent failures were within 150 m distance of the first, and practically none after 200 m (Gangl, 2008). If the geographic location of pipes is available, (Fuchs-Hanusch et al., 2012) and (Poulton et al., 2007) indicate ways to homogenize pipe lengths. In our case, GIS data were not provided, leading us to leave, merge, or split pipes dependent on their length, material and date of laying (Appendix A).

5.2.2 Pipe failure and replacement model

The used probabilistic Weibull-exponential pipe failure model is described in Scheidegger et al. (2013). It models the time between the first failure and the laying date t_0 (in years) with a Weibull distribution with shape parameter θ_1 and scale parameter θ_2 so that

$$p_1(t|t_0,\Theta) = \frac{\theta_1}{\theta_2}\left(\frac{t-t_0}{\theta_2}\right)^{\theta_1-1} e^{-\left(\frac{t-t_0}{\theta_2}\right)^{\theta_1}} \tag{31}$$

and the times between subsequent failures as exponential distributions with scale parameter θ_3:

$$p_i(t|t_0,\dots t_{i-1},\Theta) = \frac{1}{\theta_3} e^{-\left(\frac{t-t_{i-1}}{\theta_3}\right)} \quad ,i > 1 \tag{32}$$

where t_i denotes the point in time of the ith failure. To consider m different pipe characteristics m-1 regression coefficients $\beta_1 \dots \beta_{m-1}$ are estimated together with Θ. The parameter vector for pipe k is then computed as

$$\Theta_k = (\theta_1, \alpha_k\theta_2, \alpha_k\theta_3)^T \tag{33}$$

where

$$\alpha_k = \beta_1^{z_{k,1}} * \dots * \beta_{m-1}^{z_{k,m-1}}$$

The indicator variables $z_{k,j}$ equal to one if the jth characteristic is met by pipe k and otherwise zero.

To estimate the failure model parameters, the influence of past replacement on the recorded data needs to be considered. To enable an unbiased estimation of these parameters, the failure model is coupled with a replacement model in which the probability π of a pipe not to be replaced after occurrence of each failure is assumed to be constant (Scheidegger et al., 2013). Replacement due to other reasons than pipe condition, i.e. managerial replacement due to collaboration with other infrastructure providers, is not covered as it has no influence on the parameter estimation and cancels out algebraically.

Model calibration

Because the data of D do not suffice to calibrate the model using purely data-driven methods such as Maximum Likelihood Estimation (MLE) (Harrell, 2001), the failure and replacement model parameters are determined by Bayesian inference. This is widely used in statistical and engineering science and has already been applied to pipe failure models (Dridi et al., 2009; Economou et al., 2009; Watson et al., 2004). Using Bayes' theorem, a prior probability distribution of the failure model parameters is updated with observed data of target water supplier D (for the concept see e.g. Gelman et al. 2004).

Estimation of prior parameter distribution

A prior distribution provides a mathematical description of the current knowledge about the parameters in question. An informative prior can be obtained by e.g. expert elicitation (the assessment of unknown quantities from experts), literature study, or analysis of additional data. Based on experience with expert elicitation for a much simpler model (Scholten et al., 2013), we judged elicitation to be considerably more complex than maximum likelihood estimation (MLE) from available data. The prior parameter distribution for utility D (61 km) was then estimated from data of three large to mid-size Swiss water utilities A-C (> 220 km distribution network each):

First, the model parameters for each network are separately determined using MLE. For each water utility u, the parameters $\Theta_u^* = \ln(\Theta_u)$ are approximately multivariate normal distributed: $p_u(\Theta^*|\mu_u, \Sigma_u)$. The parameters of the failure model Θ_u for each utility are thus lognormal distributed with $p_u(\Theta|\mu_u, \Sigma_u)$. Second, the three parameter distributions are aggregated into one prior distribution by an equally weighted mixture of distributions and smoothing to ensure unimodality (Scholten et al., 2013).

Owed to strong correlation with the other model parameters, and identifiability issues during pre-tests, π is not directly estimated for B and C. Instead, it is fixed to a defined level and the other parameters are inferred freely. To propagate the uncertainty linked to the choice of π, we assume a beta distribution with parameters α=15 and β=2.5, $\pi \sim$Beta(α,β), and perform MLE at the 0.01, 0.1, 0.2, …, 0.9, 0.99 quantiles. α and β are chosen based on expert information from water supplier B and C who estimated the probability not to be replaced after a failure (π) as approx. 0.88-0.82 (B) and 0.88-0.97 (C) for the last 1-3 years. The resulting parameter distributions are aggregated using the probability density at the quantiles as weights to obtain one separate distribution for each B and C. Since no FC pipes are present in B and C, the same correlation to the other parameters as in network A is assumed.

Estimation of posterior parameters

The Bayesian posterior is obtained by Markov Chain Monte Carlo (MCMC) sampling using the aggregated prior of A-C, the conditional likelihood, and the network and failure data of D. Of 50'000 samples, the first 25'000 are discarded as burn-in and the posterior parameter distribution is obtained from the remaining.

Prediction of unrecorded failures

Taking the failure order as indicator of pipe condition, knowledge about the previous number of failures is needed to correctly apply condition-dependent rehabilitation strategies. Since only the times and orders of failures within the observation period are known, the number of previous failures of each pipe before the start of observations can be predicted, see supporting material A.

Prediction of future failures

Failures are predicted by embedding the failure model into the asset management software FAST (Fichtner Asset Services & Technologies, 2013). As compromise between computational time and stability, 1'000 parameter combinations randomly sampled from the posterior are imported to propagate the uncertainty of the failure model parameters. For PE pipes, further assumptions of failure model parameters are necessary given the absence of failure data for inference. The mean parameters of the Weibull distribution are set at $\theta_{1,PE}$=4.11, $\theta_{2,PE}$= 74.4 with standard deviations as $\sigma_{1,PE}$= 1.21, $\sigma_{2,PE}$= 26.73 (Scholten et al., 2013, Table 4), and $\theta_{3,PE}$= 39.7 and $\sigma_{3,PE}$= 12.8 for the exponential distribution (mean expected value; mean standard deviation of posterior θ_3 for remaining materials). After prediction and assignment of unrecorded failures to single pipes, π is no longer needed for prediction of future failures because the probability of future replacement is determined by the rehabilitation strategy.

5.2.3 Network rehabilitation model

Rehabilitation modeling in FAST is based on a system of coupled (non-linear) differential equations which describe the condition of the assets over time. Within each *aging chain* (Sterman, 2000), pipe condition is defined by the number of occurred failures governed by an age-dependent deterioration process (pipe failure model). We defined six condition classes from "zero" to "five or more" failures (Figure 12). Each pipe group is associated to its own, unique aging chain. Fifteen aging chains were implemented to model network expansion and deterioration of five pipe groups (DI1, DI2, GI3, FC, and PE), subdivided into three diameter classes (low, medium, and high criticality, section 5.2.5). Other processes that influence pipe condition over time are also modeled: network expansion, deterioration, repair, and replacement (Figure 12).

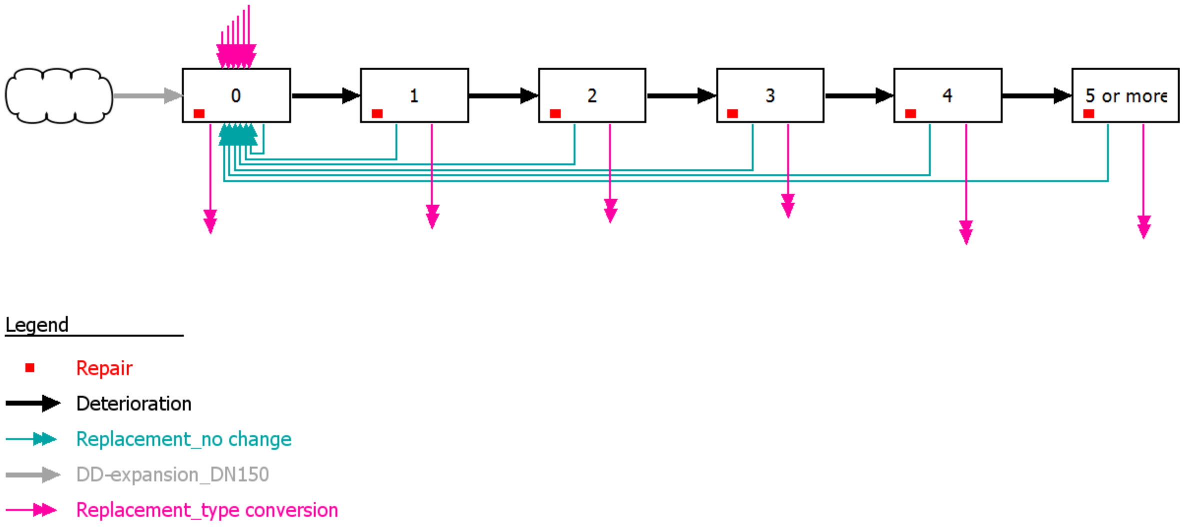


Figure 12: Exemplary aging chain with relevant processes as displayed in FAST. Boxes represent the condition state (number of failures) of its pipe members, arrows the transition between condition states and pipe groups. *DD-expansion_DN150*: distribution network expansion of 150 mm pipes; *replacement_type conversion*: replacement through pipes of another material.

Deterioration

In accord with the failure model of Scheidegger et al. (2013), the age-dependent transition from no failures to condition 1 (1st failure) is described by a Weibull distribution. The time to subsequent failures follows an exponential distribution with identical parameters. Scheidegger et al. (2013) made this choice based on the manageable complexity of this model layout and its successful application in the past by Mailhot et al. (2000).

Reactive rehabilitation (repair)

To warrant continuous water supply, we assume that all failed pipes are immediately repaired. Thereafter, a pipe is considered fully functional but one condition class higher (worse) on the aging chain due to the higher failure order.

Proactive rehabilitation (replacement)

A defined number of pipes with specified characteristics are replaced by new pipes (condition 0). The amount and characteristics depend on the rehabilitation strategy. Historical materials which are no longer available, i.e. DI1, GI2, GI3, and AC, are replaced by other materials used in Switzerland (PE pipes replace FCAC, DI2 replaces GI2, GI3, and DI1). Failed pipes are removed from the aging chain and an equal number of new pipes are created in the target aging chain of the same or new material. All other materials pipes are replaced by new pipes of the same material. It is also possible that pipes without failures are removed. One example is managerial replacement caused by collaborative ground works with other infrastructure providers or for other reasons requiring the removal of a specific material such as asbestos pipes. Managerial replacement is not considered in this study.

5.2.4 MCDA framework

MCDA allows exploring different *alternatives* (in engineering terms: options, measures, strategies, solutions, scenarios) regarding their performance on *fundamental objectives* (criteria, goals). The preferences of stakeholders are quantified based on *attributes* (quantitative performance indicators, metrics) associated to the objectives. The performance of an alternative is based on combining the prediction of its *outcome* (e.g. expected costs) with the preferences of the stakeholders for this outcome (Eisenführ et al., 2010; Keeney and Raiffa, 1993).

In the first structuring phase, the decision problem and boundary conditions are defined and main stakeholders identified (see Lienert et al. 2013a, b). Objectives, attributes, and alternatives are formulated. Secondly, the outcomes (attribute levels) of each alternative are predicted, e.g. from model outputs or expert estimates. Then subjective preferences of the decision makers (and other stakeholders) regarding the objectives are elicited. By help of a multi-attribute value model (MAVM), the overall value of each alternative is calculated by combining the outcomes with the individual preferences. The alternatives are ranked, based on overall values and discussed with the decision maker(s).

5.2.5 Objectives and attributes

Predominantly economic, hydraulic, water quality, and reliability criteria should be included in rehabilitation decision models (Engelhardt et al., 2000; Selvakumar and Tafuri, 2012). Most of these "criteria", however, are poorly formulated in terms of decision analysis because the fundamental objectives remain unclear, or because they more likely represent attributes (e.g. life cycle cost) or means objectives (e.g. low failure rate, good system condition). Means objectives are pursued to achieve another, more fundamental objective and indicate a poorly designed system of objectives (Eisenführ et al., 2010). A reformulation of the criteria mentioned in (Engelhardt et al., 2000; Selvakumar and Tafuri, 2012) results in at least three fundamental objectives of good rehabilitation strategies which we use to compare alternatives (but with other attributes; see also discussion of objectives and attributes in Lienert et al. 2014b):

1) low costs (mentioned: cost of replacement / damage / repair / maintenance / leakage and water loss / life cycle cost),
2) high reliability (mentioned: probability / percentage of the time the system is operational / ability to supply required quantity and quality of water),
3) high intergenerational equity (mentioned: failure / break rate / net present value [for financial sustainability]).

Low costs (attribute: % of mean annual per capita income)

Costs are expressed as percentage of the mean annual per capita income in the region (viz. 65'093 CHF in 2010) and are affected by future development (Appendix B). Only direct costs for repair and replacement are considered. Unit costs are 6'500 CHF per failure (median in neighboring utility, 2005-2010) covering repair, disinfection, and temporary above-ground services during interruption. Replacement cost is 910 CHF m^{-1}, including valves and fittings (mean rate charged by local engineering companies for open trench replacement). We use real incomes and assumptions about real income changes under the four future scenarios (section 5.2.9) and relate annual costs to annual incomes to unlink costs and inflation. The resulting percentages are then independent of any assumptions regarding future inflation and discount rates. This choice is also beneficial in view of elicitation from decision makers. It avoids an anchoring to certain absolute monetary levels compared to which higher future costs can be perceived as loss (reference point effect, see Kahneman and Tversky, 1979) even though the relative percentage compared to the mean income is the same.

High reliability (attribute: system reliability)

The reliability of a system (*R*) is linked to the frequency and impact of interruptions (Farmani et al., 2005; Mays, 1996). In the absence of detailed hydraulic models, we use a criticality index *C* to represent the severity of a failed pipe's impact. Assuming that larger pipe diameters result in higher property damage and number of people affected (at least in ramification networks as typical for small networks), pipes are rated into three criticality classes depending on inner diameter. Small distribution pipes (usually

≤ 150 mm): C_{low} = 1, intermediate distribution pipes (150-250 mm): C_{medium}= 5, major distribution pipes and trunk mains (≥ 250 mm): C_{high} = 10.

$$R = 1 - \frac{\sum_{i=1}^{3} C_i \cdot n_{f,i}}{\sum_{i=1}^{3} C_i \cdot n_i} \tag{34}$$

with C_i… criticality index (or importance weight) of diameter group
$n_{f,i}$…number of pipe failures in diameter group
n_i… number of all pipes in diameter group

High intergenerational equity (attribute: degree of rehabilitation)

The mean failure rate (failures per km and year) of an alternative compared to a reference (no replacement) indicates the degree of implementation of the rehabilitation demand D_{reha}, or "degree of rehabilitation".

$$D_{reha} = 1 - \frac{r_s}{r_{ref}} \tag{35}$$

with r_s… failure rate of strategic alternative s (failures per km and year)
r_{ref}… failure rate of reference strategy A_{ref} (failures per km and year)

If the rehabilitation demand of a generation is not responded to, the average age of the network and its likelihood of failure, water losses, and water quality impairment increases. Consequentially, future generations have to invest potentially higher efforts than needed by the current generation to maintain a good condition.

Uncertainty of attribute predictions

The uncertainty of the attribute predictions results from the failure predictions. These predictions incorporate the random behavior of pipe failures and the uncertainty due to parameter uncertainty of the model described in section 5.2.2. Variation under the four different future scenarios arises from the parameters assumed for network expansion and socio-economic development (section 5.2.9). Further plots regarding the sensitivity of the attribute outcomes to different criticality indices and unit costs are shown in the supporting information (section C6).

5.2.6 Strategic rehabilitation alternatives

We compare 18 strategic rehabilitation alternatives which follow three qualitative regimes: *minimal*, *average*, and *extensive* (Table 1). Failures are always repaired, regardless of the alternative. *Minimal* stands for mostly reactive alternatives, i.e. only pipes of very bad condition are replaced, a common strategy in many places (Selvakumar and Tafuri, 2012). The *average* regime describes simple replacement strategies of moderate effort, e.g. reaching a predefined lifespan or a certain number of failures (e.g. 3rd, 4th) . The *extensive* regime contains more elaborate strategies typical for large water utilities. Performance is assessed over 40 years, until 2050. To understand long-term outcomes over more than one pipe generation, calculations are done until 2110.

Table 18: Strategic rehabilitation alternatives. Failures are repaired in all alternatives. The strategies are not adapted over time, i.e. if all pipes in the worst condition states (e.g. 5 or more failures) are replaced, pipes from the next-worst condition class (e.g. 4, 3 and so on) are replaced. If there are more pipes in a certain condition class of an aging chain than should be replaced (e.g. 20 pipes in worst condition, but only 2 are replaced), the oldest pipes are selected.

Alternative		#	Description	Regime
Reference	A_{ref}	1	no. of failures if only repairs are done. i.e. function is maintained but condition deteriorating	none
Based on no. of failures (condition)	$A_{f2\ \ 5+}$		replacement only if a certain condition, applies:	
		2	- A_{f2+}: replacement after 2nd failure	average
		3	- A_{f3+}: replacement after 3rd failure	
		4	- A_{f4+}: replacement after 4th failure	minimal
		5	- A_{f5}+: replacement after 5th failure	
	$A_{f0.5\%...2\%}$		% of network replaced by condition: worst condition first*	
		6	- $A_{f0.5\%}$: 0.5 % of network	average
		7	- $A_{f1\%}$: 1 % of network	
		8	- $A_{f1.5\%}$: 1.5 % of network	extensive
		9	- $A_{f2\%}$: 2 % of network	
Based on pipe age	$A_{cyc80\ \ 100}$		all pipes older than defined replacement cycle are replaced	
		10	- A_{cyc100}: replacement cycle = 100 years	average
		11	- A_{cyc80}: replacement cycle = 80 years	
	$A_{a0.5\%...2\%}$		% replacement by age, eldest first	
		12	- $A_{a0.5\%}$: 0.5 % of network	average
		13	- $A_{f1\%}$: 1 % of network	
		14	- $A_{a1.5\%}$: 1.5 % of network	extensive
		15	- $A_{a2\%}$: 2 % of network	
Based on no. of failures and risk (pipe criticality)	$A_{fr1\%...2\%}$		% replacement by condition, riskiest first*	
		16	- $A_{fr1\%}$: 1 % of network	extensive
		17	- $A_{fr1.5\%}$: 1.5 % of network	
		18	- $A_{fr2\%}$: 2 % of network	

5.2.7 Modeling preferences

In the MCDA, "objective" outcomes of each alternative (e.g. the total costs) are combined with the "subjective" preferences of the decision maker into an overall value (see e.g. Eisenführ et al., 2010). To be able to compare very different types of attributes (e.g. costs with system reliability) on equal footing, the attribute levels are converted to a neutral value between and including 0 and 1 with help of a value function v(x). For each alternative *A*, the different values (outcomes) of each attribute are aggregated to derive the overall value V(*A*). For the aggregation, weights are needed, which reflect the relative importance that the decision maker assigns to the different attributes (or objectives). Hence, following components of the multi-attribute value model describe specific aspects of the decision makers' preferences:

Weights w_j (scaling factors) represent the relative importance of an objective j to the other objectives conditional on the range of possible attribute levels xj and take values within [0,1]. If an additive aggregation model is used, the weights sum up to 1.

Single-attribute (or marginal) value functions $v_j(x_j)$ describe how well objective j is fulfilled by achieving attribute levels x_j, thus converting attribute levels to dimensionless values between 0 (worst level, e.g. highest expected costs) to 1 (best level; lowest expected costs). Measurable value functions not only order, but also allow for strength of preference statements (Dyer and Sarin, 1979). Here, we use a common function, the exponential (measurable) value function.

$$v_j(x_j) = \begin{cases} \dfrac{1 - e^{-c_j \tilde{x}_j}}{1 - e^{-c_j}}, & c \neq 0 \\ \tilde{x}_j, & c = 0 \end{cases} \tag{36}$$

with $\tilde{x}_j = (x_j - \min(x))/(\max(x) - \min(x))$. Constant c_j determines whether the function is concave (> 0), convex (< 0) or linear (= 0). The value functions are defined over the range of the alternatives' outcomes, rounding up resp. down to the nearest 0.05 multiple for the degree of rehabilitation and 0.01 for reliability and costs.

A multi-attribute aggregation function aggregates the preference information of weights assigned to the different objectives and the values achieved for each attribute into one score returned from the MAVM, the overall value V(A) ∈ [0,1] of each alternative A. An overall value of 1 means that the outcomes of an alternative regarding all objectives are on their best level (i.e. here: costs are on their lowest-possible level, system reliability and degree of rehabilitation on their highest-possible level). Because of its simplicity, the additive model is often used (Eisenführ et al., 2010). The overall additive value of alternative A is

$$V(A) = \sum_{j=1}^{m} w_j \cdot v_j(x_j(A)) \quad ; \quad \sum_{j=1}^{m} w_j = 1 \tag{37}$$

and the additive weights sum to unity. Value functions describe preferences under certainty. For risky (uncertain) outcomes, multi-attribute utility functions (Keeney and Raiffa, 1993) are required, with additional axioms to be satisfied. Value functions can be transformed into utility functions if the decision maker's intrinsic risk attitude is known (Dyer and Sarin, 1982; Keeney and Raiffa, 1993). For risk neutral decision makers, value and utility functions coincide.

Table 19: Preference parameters for local sensitivity analysis (reliab= reliability, reha= intergenerational equity). 1st set: sensitivity of different weights attributed to the three objectives, assuming linear value functions. 2nd set: sensitivity to different shapes of value functions, assuming equal weights.

	preference	w_1 (reliab)	w_2 (costs)	w_3 (reha)	c_1 (reliab)	c_2 (costs)	c_3 (reha)
weights	v.lin.eqw	1/3	1/3	1/3	0.00	0.00	0.00
	v.lin.w1a	1.00	0.00	0.00	0.00	0.00	0.00
	v.lin.w2a	0.00	1.00	0.00	0.00	0.00	0.00
	v.lin.w3a	0.00	0.00	1.00	0.00	0.00	0.00
	v.lin.w1h	0.50	0.25	0.25	0.00	0.00	0.00
	v.lin.w2h	0.25	0.50	0.25	0.00	0.00	0.00
	v.lin.w3h	0.25	0.25	0.50	0.00	0.00	0.00
v(x)	v.1cv.eqw	1/3	1/3	1/3	-4.00	0.00	0.00
	v.2cv.eqw	1/3	1/3	1/3	0.00	-4.00	0.00
	v.3cv.eqw	1/3	1/3	1/3	0.00	0.00	-4.00
	v.acv.eqw	1/3	1/3	1/3	-4.00	-4.00	-4.00
	v.1cc.eqw	1/3	1/3	1/3	4.00	0.00	0.00
	v.2cc.eqw	1/3	1/3	1/3	0.00	4.00	0.00
	v.3cc.eqw	1/3	1/3	1/3	0.00	0.00	4.00
	v.acc.eqw	1/3	1/3	1/3	4.00	4.00	4.00

For simplification, we assume that there is only one decision maker. In a real decision situation, the parameters of the MAVM are typically inferred from preference statements of each stakeholder separately (methods for elicitation of the weights, value/utility functions, and aggregation function are presented in e.g. Eisenführ et al., 2010; Keeney and Raiffa, 1993). We assess the influence of different preferences on the alternative ranking with a local sensitivity analysis over varying weights and value functions (Table 19).

5.2.8 Dominance and ranking of alternatives under uncertainty

To reduce unnecessary complexity in MCDA, it is recommended to exploit dominance relationships as first step (e.g. Eisenführ et al., 2010). Hereby, the analysis is simplified by removing dominated (hence irrelevant) alternatives before calculating the overall values (or utilities). For risky outcomes, stochastic dominance concepts can be used (Hadar and Russell, 1969; Hanoch and Levy, 1969; Rothschild and Stiglitz, 1970).

First- degree stochastic dominance (FSD) is fulfilled if alternative A's probability of achieving better attribute levels than alternative B is higher for at least one attribute and equally high for all others. FSD can be determined graphically using risk profiles 1-P(X) of the attributes' cumulative probability functions P(X) (Eisenführ et al., 2010). A dominates B regarding attribute x if the risk profile of A is always above that of B. If the risk profiles intersect, additional information about the decision makers' preference under risk is needed to determine dominance. Practically, for each year between 2010 to 2050, the outcome of the three attributes for each of the 1000 parameter samples are computed. From these results, the cumulative probabilities are calculated.

For risk averse decision makers, second-degree stochastic dominance (SSD) delivers further insights. SSD is satisfied if the area under the cumulative probability curve of B exceeds the cumulated area under that of A for all x (Graves and Ringuest, 2009). As the necessary pairwise comparisons of distributions get computationally very expensive for 18 alternatives under four scenarios, we use the mean and risk-adjusted mean-Gini summary statistic (Graves and Ringuest, 2009). In the mean-Gini model, mean μ and risk-adjusted mean μ' (Gini's Mean Difference, GMD) of the alternatives are compared directly (Shalit and Yitzhaki, 1994). the mean attribute outcome of *A* is larger than or equal to that of *B*, $\mu_A \geq \mu_B$, and if

$$\begin{gathered}\mu'_A \geq \mu'_B \text{ or} \\ \mu_A - 2\,cov\big(X_A, P_A(X_A)\big) \geq \mu_B - 2\,cov\big(X_B, P_B(X_B)\big)\end{gathered} \tag{38}$$

where X_A is the random variable describing the attribute outcome of alternative *A*, and $P_A(X_A)$ is its cumulative distribution, see (Yitzhaki, 2003). Conveniently, this approach is not only applicable to non-normal probability distributions, but also fulfills the necessary conditions of SSD without requiring pairwise comparisons. If the risk profiles cross once at most, the sufficient conditions for SSD are additionally fulfilled (Shalit and Yitzhaki, 1994). Practically, alternatives are ranked by μ and μ' of the outcomes between 2010 and 2050. Those with better ranks dominate those with worse ranks whenever the rank relationship order of μ and μ' is maintained (Graves and Ringuest, 2009). To establish an overall rank for comparison within and across scenarios during sensitivity analysis considering different preferences, the average of μ and μ' of the aggregated value (eq. 36) per alternative and set of different parameters (Table 19) is used.

5.2.9 Robustness under four future scenarios

Four future development scenarios were formulated: *Status quo* (no change / baseline), *Boom* (massive growth), *Quality of life* (qualitative growth), and *Doom* (decline). Their characteristics cover a range of technical, environmental, and socio-economic aspects, see Lienert et al. (2013b) for details and Appendix B for a summary of the information relevant to this work.

Diverging notions about robustness prevail in the decision sciences and operational research (Roy, 2010). We mean robustness in the context of stability and sensitivity, i.e. how stable the ranking of alternatives under different future scenarios is.

Following Goodwin and Wright (2001), all alternatives are separately evaluated and ranked under each future scenario. Their approach assumes that the preferences are independent of the scenario and that consequently, only the attribute outcomes depend on the scenarios. This is in contrast to the assumption of different preferences under each future scenario (Montibeller et al., 2006; Stewart et al., 2013) , where for example, the costs might be judged relatively more important in a dire economic future scenario than in a prospering future scenario. We propose to consider changing preferences due to learning and different boundary conditions as part of an adaptive management plan. Hereby validation – or if necessary – re-assessment of the decision makers' preferences after some time would be necessary. This seems less problematic than eliciting hypothetical scenario-adjusted preferences from decision makers others have resorted to (Karvetski et al., 2009b; Ram and Montibeller, 2013). In our case, the overall robustness of each alternative is derived from changes in the rankings under the four scenarios.

5.2.10 Implementation

Except rehabilitation modeling in FAST , data handling, parameter inference, preference modeling, and evaluation are implemented in the freeware language and environment for statistical computing R (R Development Core Team, 2011) and supported by R packages: *optimx* (Nash and Varadhan, 2011), *DEoptim* (Mullen et al., 2011), *adaptMCMC* (Scheidegger et al., 2011), *utility* (Reichert et al., 2013), and *ggplot2* (Wickham, 2009).

5.3 Results

5.3.1 Network data

The length distributions of the four water suppliers' raw data are strongly diverging (Figure 13). Modal pipe lengths decrease from water supplier A to D, as well as distances between the 5 to 95 % and 25 to 75 % quantiles. After homogenization, water networks A to C share similar distributional properties. The goal of creating homogeneous lengths of 100-200 m was achieved for at least 75 % of pipes in A-C, but less in D.

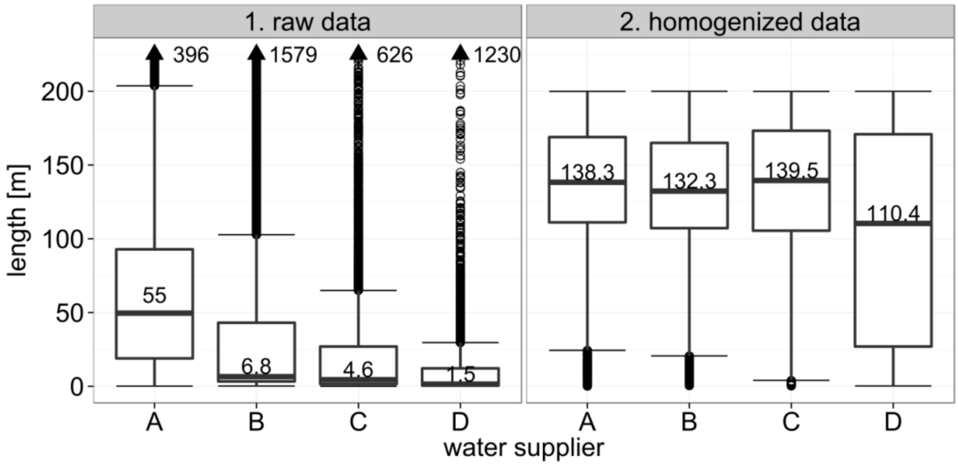


Figure 13: Pipe length distributions before and after length homogenization. The boxes and whiskers represent the 5, 25, 75, and 95 % quantiles; the thick horizontal line indicates the modal length of pipes in network A-D.

Figure 14, shows the material distributions of the four networks. The largest portions are ductile cast iron (DI1, DI2) and grey cast iron (GI2, GI3) pipes, followed by differing portions of fiber / asbestos cement (FC), steel (ST), and polyethylene pipes (PE) installed mostly after 1950.

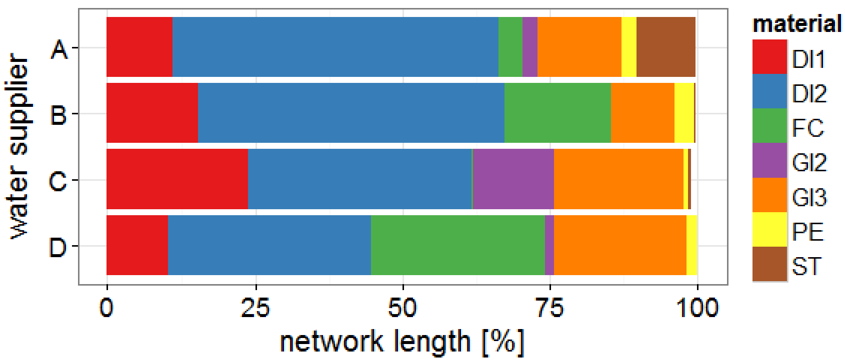


Figure 14: Material proportions in the four water supply networks. DI1 and DI2: ductile iron pipes (1st :1964-80; 2nd : > 1980), FC: fiber and asbestos cement, GI2 and GI3: grey cast iron (2nd: <1930, 3rd : >1930), PE: polyethylene, and ST: steel.

Although DI2 is the most prevalent material, only few recorded failures are available in utilities B-D (Table 20). Additionally, there are no or very few higher order failures on DI2 pipes in B-D. This can lead to parameter estimation difficulties, also for other materials with few recorded failures (FC, ST). Most failures were recorded on DI1 and GI3 pipes with proportionally more failures in network A and C, also regarding higher order failures.

Table 20: Network characteristics and failures of the four water networks (A-D) after length homogenization.

	A	B	C	D
observation period	2000-2010	2001-2011	1996-2011	2001-2010
total length [km]	715	385	227	61
ø pipe length [m]	134.7	127.3	129.2	102.0
total failures/ *higher-order failures*	669/*233*	182/*32*	279/*97*	40/*2*
DI1	140/*47*	95/*19*	89/*28*	13/*0*
DI2	133/*38*	19/*0*	12/*2*	3/*0*
GI2	46/*18*	0/0	51/*20*	0/*0*
GI3	240/*88*	59/*12*	121/*46*	18/*2*
FC	14/*0*	8/*1*	0/*0*	6/*0*
ST	96/*42*	0/*0*	1/*0*	0/*0*
PE	0/*0*	1/*0*	3/*0*	0/*0*

5.3.2 Failure model

The estimated failure model parameters from MLE (networks A-C), the aggregated prior, and the posterior parameters are presented in Table 21. Parameters from MLE with fixed π of B and C are shown in the supporting material (Table C.1). Networks A-C show the same ordering of times to failure, FC ≥ DI2 >> GI3 ≥ DI1, despite considerable differences in the parameters. This order is also maintained in the resulting prior and posterior distributions.

Table 21: Summary statistics of the marginal parameter distributions of networks A–C individually and aggregated, as well as the posterior for network D. For B and C, only the aggregated parameter distributions of eleven MLE runs each with fixed π are shown.

		all	DI1		DI2		GI3		FC	
		$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\beta}_{DI2}\hat{\theta}_2$	$\hat{\beta}_{DI2}\hat{\theta}_3$	$\hat{\beta}_{GI3}\hat{\theta}_2$	$\hat{\beta}_{GI3}\hat{\theta}_3$	$\hat{\beta}_{FC}\hat{\theta}_2$	$\hat{\beta}_{FC}\hat{\theta}_3$
Posterior (A-C,D)	$\hat{\boldsymbol{\theta}}$	***1.47***	***72.1***	***20.5***	***217.0***	***62.1***	***89.7***	***25.8***	***274.7***	***81.3***
	$\mathrm{sd}(\hat{\theta})$	*0.18*	*14.3*	*5.1*	*61.11*	*22.0*	*12.7*	*6.8*	*78.3*	*37.6*
Prior (A-C)	$\hat{\boldsymbol{\theta}}$	***1.60***	***77.1***	***17.3***	***195.7***	***44.8***	***88.7***	***20.2***	***280.3***	***70.4***
	$\mathrm{sd}(\hat{\theta})$	*0.24*	*20.0*	*6.8*	*65.7*	*22.4*	*16.1*	*8.2*	*122.8*	*55.6*
A	$\hat{\boldsymbol{\theta}}$	***1.59***	***70.0***	***10.1***	***159.9***	***23.0***	***86.8***	***12.5***	***154.0***	***22.2***
	$\mathrm{sd}(\hat{\theta})$	*0.13*	*14.3*	*2.01*	*30.0*	*4.0*	*18.6*	*2.7*	*35.7*	*5.1*
B	$\hat{\boldsymbol{\theta}}$	***1.75***	***59.5***	***22.2***	***169.8***	***63.1***	***76.4***	***28.5***	***304.3***	***113.2***
	$\mathrm{sd}(\hat{\theta})$	*0.27*	*6.1*	*4.8*	*53.3*	*22.3*	*8.9*	*6.0*	*94.5*	*40.3*
C	$\hat{\boldsymbol{\theta}}$	***1.43***	***97.2***	***16.6***	***245.7***	***41.7***	***95.7***	***16.4***	**-**	**-**
	$\mathrm{sd}(\hat{\theta})$	*0.19*	*13.8*	*2.6*	*79.6*	*12.5*	*11.9*	*2.6*	-	-

Whereas the Weibull and exponential scale parameters ($\hat{\theta}_2, \hat{\theta}_3$) of FC and DI2 are of similar magnitude in network A, in network B the parameters for FC are significantly larger. DI1 and GI3 pipes are, according to the magnitude of the parameters, most durable in network C ($\hat{\theta}_{2,DI1} = 97.2, \hat{\theta}_{2,GI3} = 95.7$), followed by A ($\hat{\theta}_{2,DI1} = 70.0, \hat{\theta}_{2,GI3} = 86.8$) and then B ($\hat{\theta}_{2,DI1} = 59.5, \hat{\theta}_{2,GI3} = 76.4$). The uncertainty of the DI2 and FC parameters is considerable in A-C, also in the aggregated prior and posteriors. As the smaller variance of the posterior indicates, something could be learned even from the (few) data of network D, especially for DI1 and GI3.

Because some pipe rehabilitation strategies are condition-based, failures before the start of formal failure recording were predicted for D (i.e. failures before 2001). The predicted number of failures is 149 and results from a single run of the prediction model as described in section 5.2.2.

5.3.3 Outcomes of strategic alternatives

The outcomes of the 18 alternatives regarding costs, reliability, and intergenerational equity over time are visualized in Figure 15. Here, we show the relative performance of each alternative for each of the three attributes alone, without considering possible preferences of decision makers and without aggregating to an overall value for each alternative in the MCDA. Note that the outcomes for reliability and intergenerational equity are identical in the Status quo and Doom scenario (because of identical framework conditions).

Compared by their median outcomes (lines), $A_{f1.5\%}$ and $A_{f2\%}$ (global replacement by condition; see Table 18; purple) and $A_{a1.5\%}$ and $A_{a2\%}$ (global replacement by age; red) often outperform the other alternatives - visible from them being below the others for costs, and above for reliability and intergenerational equity. Notably, the median outcomes of the condition-risk dependent strategies ($A_{fr1...2\%}$; blue lines) perform rather badly compared to less sophisticated alternatives (e.g. $A_{cyc80...100}$, orange; $A_{f2...5+}$; green lines). The median of the reference alternative A_{ref} (solid black line) performs worst for all attributes, except for costs in all scenarios.

Since the 0.05-0.95 inter-quantile ranges of the alternatives (shaded areas) regarding reliability and rehabilitation are large and considerably overlap, any ranking based on the attribute outcomes alone is speculative. The outcomes change substantially after the defined planning horizon 2050, such that the extension of the evaluation horizon to 2110 could potentially result in a different ranking.

Looking at costs separately, Figure 15 displays a continuous increase over time for all alternatives except $A_{cyc80...100}$ in the Doom scenario. In the other scenarios, the costs of all alternatives initially decrease and then stabilize or increase again slightly. Costs are highest in the Doom scenario, the maximum increase expected for alternatives $A_{fr2\%}$ and $A_{f2\%}$ (median costs about 0.4% in 2050, 1.1% in 2110). The median costs of other alternatives in the Doom scenario increase at lower rates, except for the cyclic alternatives (A_{cyc80}, A_{cyc100}; orange). Peak costs of the cyclic alternatives indicate peak investments (also in the other scenarios), reaching up to 7.11 % for A_{cyc100}. In the Status quo, costs for all alternatives decrease slightly and stabilize for all alternatives except $A_{cyc80...100}$.

Reliability increases strongly in the Boom and Quality of Life scenario until about 2030-2050, and especially abruptly for A_{ref} and risk-condition dependent rehabilitation alternatives (A_{ref}, $A_{fr1\%...2\%}$; blue). It stabilizes after 2050 between 1 and 0.99 or decreases slightly (A_{cyc80}, A_{cyc100}, $A_{f2...5+}$). Reasons for this abrupt change are discussed in section 5.4.3. It comes along with a strong improvement of the degree of rehabilitation until 2050 (up to 90 %) but also a strong setback, especially in the Boom scenario, with only slow recovery thereafter. In the Doom and Status quo scenarios, reliability decreases for A_{ref}, $A_{fr1...2\%}$, A_{f5+}, as well as A_{f4+} and increases for the other alternatives (until stabilization).

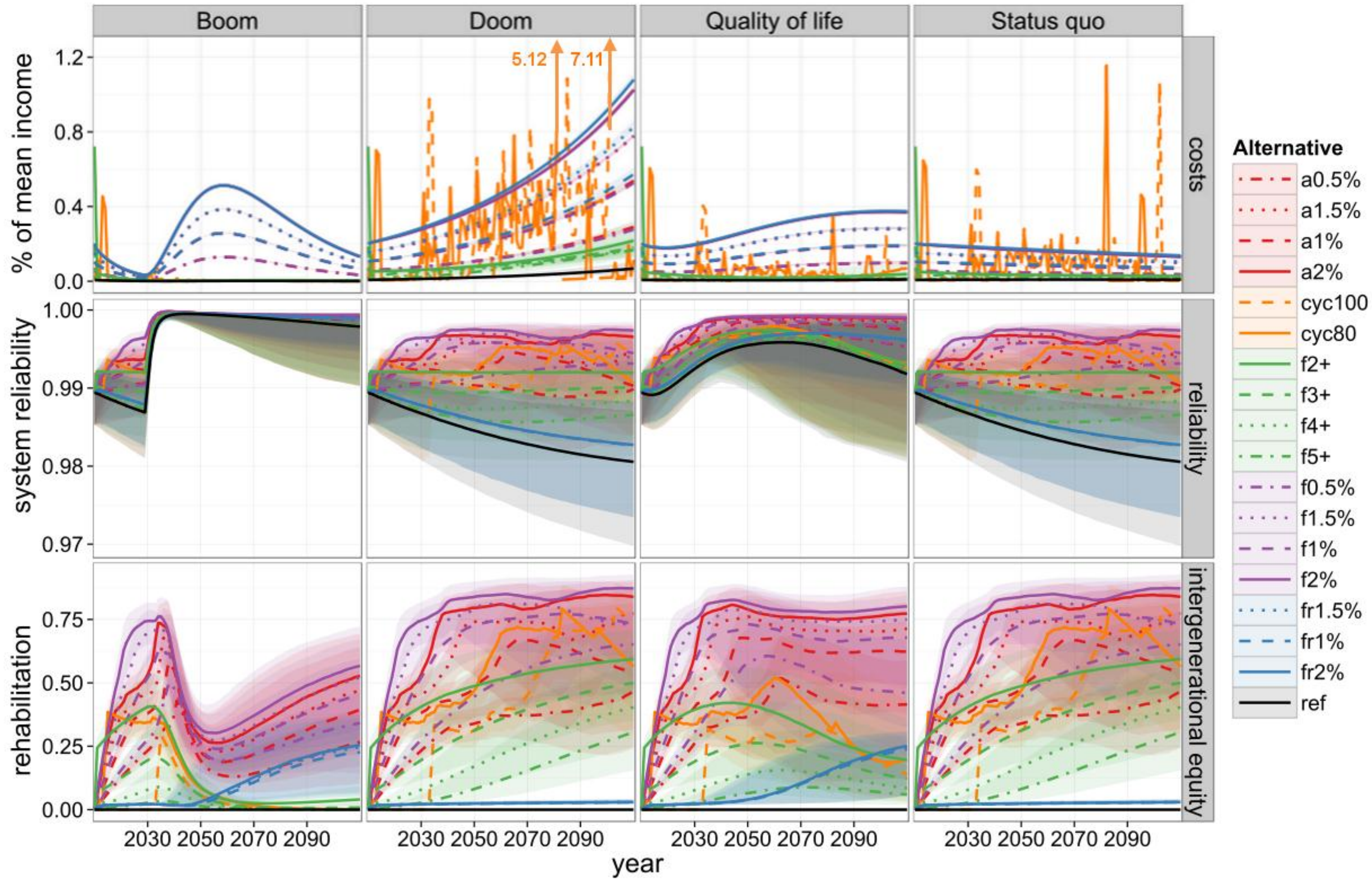


Figure 15: Outcomes of 18 strategic planning alternatives under four scenarios until 2110. We show the outcomes on the attribute levels: % of mean income, system reliability as R based on the criticality index, and rehabilitation as D_{reha} based on failure rates (see 5.2.5). These results do not contain assumptions about the preferences of decision makers, and thus there is no aggregation of the three attributes to an overall value for each alternative (as done later in the MCDA). More results can be found in the additional tables and figures of the supporting information. Lines represent the 0.5 (median), shaded areas the 0.05-0.95 quantiles. Costs improve with decreasing values, reliability and intergenerational equity with increasing values. Note that for better visibility the % mean income is zoomed in, and two peaks exceeding the visible range are indicated by arrows. Costs for $A_{a1\ \ 2\%}$ and $A_{f1\ \ 2\%}$ overlap with $A_{fr1\ \ 2\%}$ under most scenarios.

Table 22: Mean attribute ranks and risk-adjusted mean attribute ranks of 18 strategic alternatives over the time horizon 2010-2050. Shaded: dominated alternatives. Future scenarios: BO- Boom, DO- Doom, QG- Quality of Life, SQ- Status quo.

	Alternative	$A_{f2\%}$	$A_{f1.5\%}$	$A_{a2\%}$	$A_{a1.5\%}$	A_{cyc80}	A_{f2+}	$A_{f1\%}$	$A_{a1\%}$	$A_{f0.5\%}$	$A_{a0.5\%}$	A_{f3+}	$A_{fr2\%}$	$A_{fr1.5\%}$	$A_{fr1\%}$	A_{f4+}	A_{f5+}	A_{cyc100}	A_{ref}
Costs (mean annual per capita income)																			
BO	rank(μ_{cost})	16	15	17	14	7	6	10	11	8	9	5	18	13	12	4	2	3	1
	rank(μ'_{cost})	16	15	17	14	6	7	10	11	8	9	5	18	13	12	4	3	2	1
DO	rank(μ_{cost})	16	13	17	14	10	8	9	11	6	7	4	18	15	12	3	2	5	1
	rank(μ'_{cost})	16	13	17	14	6	7	10	11	8	9	5	18	15	12	4	3	2	1
QG	rank(μ_{cost})	16	13	17	14	9	6	10	11	7	8	4	18	15	12	3	2	5	1
	rank(μ'_{cost})	16	13	17	14	6	7	10	11	8	9	5	18	15	12	4	3	2	1
SQ	rank(μ_{cost})	16	13	17	14	9	8	10	11	6	7	4	18	15	12	3	2	5	1
	rank(μ'_{cost})	16	13	17	14	6	7	10	11	8	9	5	18	15	12	4	3	2	1
Reliability (system reliability)																			
BO	rank($\mu_{reliab.}$)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	rank($\mu'_{reliab.}$)	1	2	3	4	6	8	5	7	9	10	11	12	13	14	15	17	16	18
DO	rank($\mu_{reliab.}$)	1	2	3	4	7	6	5	8	9	10	11	14	15	16	13	17	12	18
	rank($\mu'_{reliab.}$)	1	2	3	5	6	8	4	7	9	10	11	14	15	16	13	17	12	18
QG	rank($\mu_{reliab.}$)	1	2	3	4	8	7	5	6	9	10	11	12	13	15	16	17	14	18
	rank($\mu'_{reliab.}$)	1	2	3	4	7	8	5	6	9	10	11	13	14	15	16	17	12	18
SQ	rank($\mu_{reliab.}$)	1	2	3	4	7	6	5	8	9	10	11	14	15	16	13	17	12	18
	rank($\mu'_{reliab.}$)	1	2	3	5	6	8	4	7	9	10	11	14	15	16	13	17	12	18
Intergenerational equity (degree of rehabilitation)																			
BO	rank($\mu_{rehab.}$)	1	2	3	4	8	7	5	6	9	10	11	14	15	17	13	16	12	18
	rank($\mu'_{rehab.}$)	1	2	3	5	9	7	4	6	8	10	11	14	15	17	13	16	12	18
DO	rank($\mu_{rehab.}$)	1	2	3	5	7	6	4	8	9	10	11	15	16	17	13	14	12	18
	rank($\mu'_{rehab.}$)	1	2	3	5	8	6	4	7	9	10	11	15	16	17	13	14	12	18
QG	rank($\mu_{rehab.}$)	1	2	3	5	8	7	4	6	9	10	11	15	16	17	13	14	12	18
	rank($\mu'_{rehab.}$)	1	2	3	5	9	7	4	6	8	10	11	15	16	17	13	14	12	18
SQ	rank($\mu_{rehab.}$)	1	2	3	5	7	6	4	8	9	10	11	15	16	17	13	14	12	18
	rank($\mu'_{rehab.}$)	1	2	3	5	8	6	4	7	9	10	11	15	16	17	13	14	12	18

5.3.4 Outcomes of strategic alternatives and dominance

There is a visible ordering of risk profiles within strategy groups, indicating first-degree stochastic dominance (FSD) of some alternatives and attributes (Figure 18-20, Appendix); e.g. $A_{f2\%}$ is always better than $A_{f1.5\%}$, $A_{f1\%}$, and $A_{f0.5\%}$. This ranking is reversed regarding the cost attribute. In addition, some of the risk profiles cross (e.g. $A_{cyc80...100}$, $A_{f2...5+}$), and no clear ordering is apparent. Thus, no FSD dominance which is stable across all scenarios and attributes can be determined.

Assuming risk-aversion, the results from mean-Gini analysis are more insightful (see Table 22 for ranks, Table C.2 – C.4 in supporting material for outcomes). There is a stable dominance order for reliability and intergenerational equity regarding both mean and risk adjusted mean in the $A_{f0.5...2\%}$, $A_{a0.5...2\%}$, $A_{cyc80...100}$, and $A_{f2...5+}$ groups under all scenarios. Additionally, $A_{f2\%}$ has rank 1 (best) and A_{ref} rank 18 (worst) for both attributes under all scenarios.

For costs, the rank order within groups is inversed; A_{ref} has the first rank, and $A_{f2\%}$ rank 16 under all scenarios. Nonetheless, same dominance relationships which are stable across scenarios are apparent: the mean and risk-adjusted mean of, A_{f2+} and A_{f3+} are better than those of $A_{fr1...2\%}$ under all scenarios, indicating dominance. $A_{fr1...2\%}$ are hence removed, because they will always be less preferred by a rational decision maker. Furthermore, $A_{f0.5\%}$ dominates $A_{a0.5\%}$, $A_{f1\%}$ dominates $A_{a1\%}$, and $A_{f2\%}$ dominates $A_{a2\%}$, leading to the exclusion of $A_{a0.5\%}$, $A_{a1\%}$, and $A_{A2\%}$. Finally, twelve non-dominated alternatives remain: $A_{f2\%...0.5\%}$, $A_{a1.5\%}$, $A_{cyc80...100}$, $A_{f2+...5+}$ and A_{ref}. In continuation, only these are considered.

5.3.5 Ranking and sensitivity under different preference assumptions

The ranking of the non-dominated alternatives is sensitive to alterations of the preference model, especially the weights (Figure 16), but also the value function form (Figure 17), see also Eq. 35 and 36, and Table 19. The observed rank order under the assumption of linear value functions and equal weights (V.lin.eqw, black diamond) is: $A_{f2\%} > A_{f1.5\%} > A_{f2+} \sim A_{f1\%} > A_{a1.5\%} > A_{cyc80} > A_{f0.5\%} > A_{f3+} > A_{f4+} \sim A_{cyc100} > A_{f5+} > A_{ref}$ (“Rank” in Fig. 16-17 meaning the mean rank of μ and μ’, alternatives from best to worst). The rank order of the best and worst-ranked three alternatives is inverted under all scenarios, if only costs are important (V.lin.w2a, purple squares, receiving all the weight), and also very sensitive to zero weights for intergenerational equity (V.lin.w3n, green triangles). If costs receive half the weight ($w_2 = 0.5$, V.linw2h, purple circle), only the order of the top-ranked alternatives is affected, either A_{f2+}or $A_{f1.5\%}$ becoming best- ranked and $A_{f2\%}$ third.

The ranking is less sensitive to the value function form, see Figure 17. Most distinct are the ranking changes due to all- convex value functions (V.acv.eqw, black dots), resulting in considerably worse ranks for $A_{a1.5\%}$ in all scenarios, and for $A_{f1...2\%}$ in the Boom scenario. In addition, the ranks of A_{ref}, $A_{f3...5+}$, and A_{cyc100} improve greatly. Furthermore, if only the costs value function is concave (V.2cv.eqw, blue dots), A_{f2+} becomes the best-ranked alternative while $A_{f2\%}$ and $A_{f1.5\%}$ are second to fourth-ranked. Apart from these cases, the ranking is fairly robust across scenarios and preferences.

The complete ranking and corresponding values of all alternatives without assuming risk-aversion for second-degree stochastic dominance (SSD) is shown in the supporting material (Figure C.1, C.2).

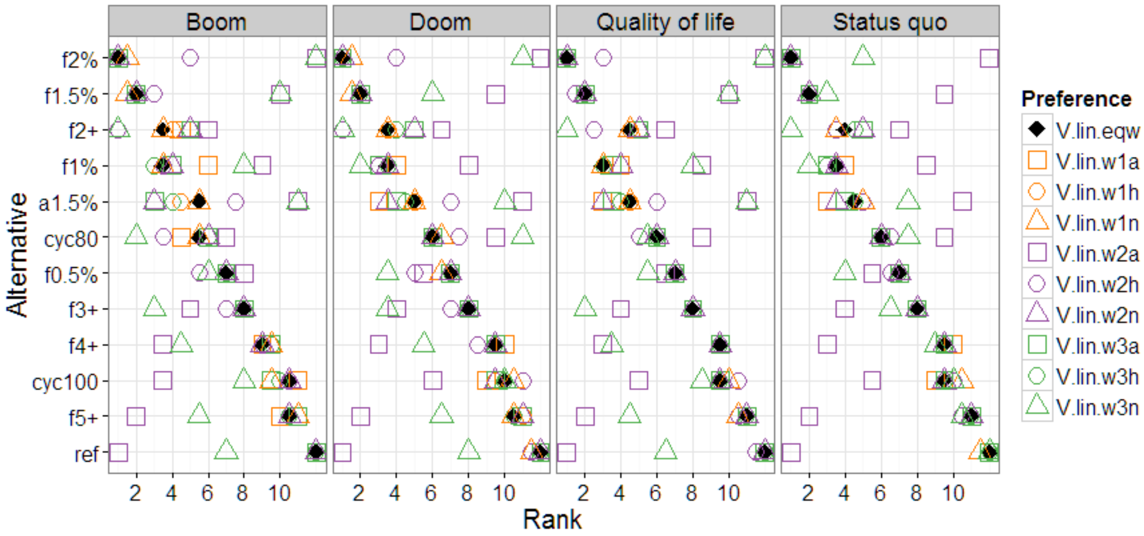


Figure 16: Sensitivity of the ranking of alternatives to weight changes under four scenarios over the time horizon 2010-2050. w_1= reliability, w_2= costs, w_3= intergenerational equity, see Table 19.

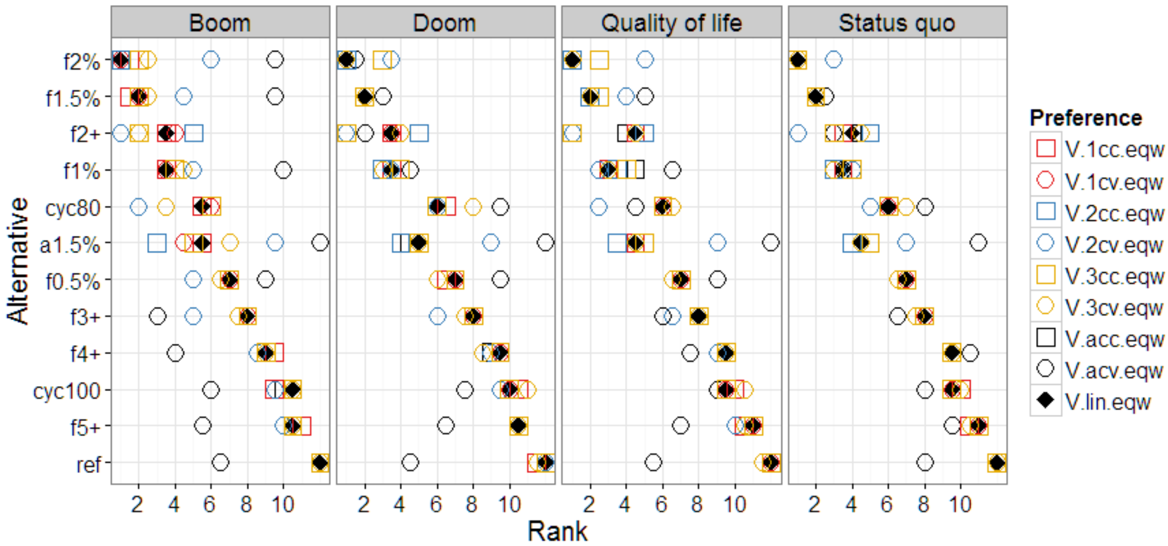


Figure 17: Sensitivity of the alternatives' ranking to value function changes under four scenarios over the time horizon 2010-2050. c_1= reliability, c_2= cost, c_3= intergenerational equity, see Table 19.

5.4 Discussion

5.4.1 Data preparation

The homogenization approach led to satisfactory homogenization of the pipe length distributions of water networks A-D, being slightly less satisfactory in the smaller pipe network D. Although more homogeneous than the raw data, many short pipes remained unmerged; likely impeded by their unique material-diameter-laying date combinations. A drawback of the approach is that merged pipes do not necessarily have a distinctive location because pipes are merged by grouping without consideration of their detailed location, see section 5.2.1. This could be improved by a GIS-based merging procedure which considers the location and other pipe characteristics (Fuchs-Hanusch et al., 2012). If electronic GIS data are unavailable, the presented novel data preparation approach delivers satisfying results for strategic asset management and the individual length of pipe sections can be overcome to reduce the influence of pipe lengths on pipe failure behavior. For tactical and operational asset management, however, the knowledge of pipe location and its consideration during pipe grouping is central both to homogenize the data accordingly and to prioritize pipe rehabilitation projects.

5.4.2 Failure and rehabilitation model

The selected failure model of Scheidegger et al. (2013) is a choice of suitability, not of conviction. Despite being reasonably simple, its big advantage is its capability of handling left-truncated and right-censored data subject to potential survival bias from deleted historical records. Together with the Bayesian approach, this makes the model suitable also for small networks.

Sensible failure model parameters for water utility D could be determined. The order of times to failure of the pipe groups (FC> DI2> GI3> DI1) is in line with results from a former analysis of pipe lifetimes in Switzerland (Scholten et al., 2013). Differences between prior and posterior parameters are visible, but small. Consequentially, the uncertainty of the failure model parameters is large which is reflected in the considerable uncertainty of the resulting attribute predictions. This is not surprising, considering the small number of observed failures (40). Consequentially, the priors (based on 1130 failures of utility A-C) are very influential. The mean parameters of material groups with few first and subsequent failures (DI2 and FC in network B, C) are remarkably large and highly uncertain. This might be indicative of lacking identifiability under purely data-driven MLE, as also observed concerning the already remediated parameter estimation with fixed π for B and C. These difficulties did not arise, however, in network A with more network and failure data. To achieve a better adaptation to local pipe failure behavior and reduce parameter uncertainty, the model parameters should be updated once additional failure data of D become available. Model validation as commonly performed with help of hold-out samples (e.g. Renaud et al., 2012) is difficult in situations where purely data-driven approaches do not suffice to parameterize the model, as mainly the consistency of the prior distributions would be tested. The use of simulated data to testify general model suitability is thus recommended (Scheidegger et al., 2011; Scheidegger and Maurer, 2012; Scheidegger et al., 2013). Formulation of the prior should be done with great care, e.g. by eliciting and discussing these with local experts (Scholten et al., 2013).

Considering that water suppliers A–C are amongst the larger and rather well-documented water networks in Switzerland, the applicability of more complex failure models applying purely frequentist inference procedures to small networks is questionable. Model simplicity, however, was traded against strong assumptions:

a. Weibull model for time to first failure: the hazard rate begins at zero, not accounting for initial failures on the "bathtub curve" (Kleiner and Rajani, 2001). Practically, this was handled by removing failures in the pipe laying year.

b. Subsequent failures are described by identical exponential distributions and therefore do not account for decreasing times between failures with increasing failure orders.

c. One covariate β_k per material used to scale both θ_2 and θ_3 does not allow for separate adjustment of time to first failure and subsequent failures relative to the baseline.

Network size and data allowing, the model of (Le Gat, 2009; Renaud et al., 2011) could be an alternative as it is based on different assumptions and also able to deal with selective survival and left-truncated-right-censored data.

Additional to future uncertainty (captured by four scenarios) failure model parameter uncertainty is propagated to the rehabilitation model outcomes. The propagation of the uncertainty adherent to the prediction of previous failures (before recording) is limited for practical reasons. Because the FAST rehabilitation model runs on one specific network of pipes with corresponding condition at a time, propagation of prediction uncertainty regarding unrecorded previous failures was impracticable. This effect is reduced by the prediction of the number of unrecorded failures prior to failure recording for each individual pipe, see section 5.2.2. If there are many pipes in the network, the overall number and distribution of previous failures over the network approximates the distribution obtained if this uncertainty was explicitly accounted for. To improve predictions for small networks, the adaptation of the software to allow for consideration of uncertainty regarding the number of failures is necessary.

5.4.3 Outcomes of strategic planning alternatives

We found that infrastructure costs (relative to the mean taxable income) increase strongly in the Doom scenario, but are rather stable, if not decreasing, in the other scenarios (Figure 15). The higher costs in the Doom scenario are due to decreasing population size and decreasing real incomes. On the contrary, the initial cost decrease in the growth scenarios (Boom, Quality of Life) can be attributed to population growth, which reduces per capita costs. Unless choosing A_{cyc80} and A_{cyc100} , peak costs arising from a group of pipes suddenly needing replacement are not likely to occur. The comparatively small uncertainty of costs (Fig.15) is due to the little influence of the uncertainty of the number of failures in light of about fifteen times higher replacement costs.

Reliability and intergenerational equity increased for most alternatives and scenarios (Figure 15). Two outcomes are surprising: 1) the strong increase in reliability and intergenerational equity under the Boom scenario until 2030 followed by a strong decrease until 2050 (less pronounced in Quality of Life), and 2) the comparatively bad performance of the condition-risk based alternatives $A_{fr1\%...2\%}$. Both can be explained by network expansion and the link to the failure rate (see also Figure C.3, supporting material). Besides improvement of pipe condition caused by the rehabilitation strategy, expansion with new pipes leads to an additional enhancement of the overall network condition. This is especially remarkable in the Boom scenario, since here, the proportion of large pipes in the network increases faster and the number of pipes per inhabitant decreases. The influence of network expansion leads even the reference alternative A_{ref} to experience a strong increase in reliability in the Boom scenario. The low performance of strategies $A_{fr1\%...2\%}$ (1 % to 2 % annual condition-based replacement by criticality), can be explained by the low number (34) of high criticality pipes in the small utility D. These strategies are more effective when there are substantial numbers of high criticality pipes in higher condition classes, as indicated by the increase in rehabilitation performance after 2050 in the growth scenarios. Additionally, their performance might improve considerably if damage costs were comprised (expecting higher damage from high-criticality pipes).

5.4.4 Ranking of alternatives and sensitivity

First-degree stochastic dominance analysis of the risk profiles did not lead to finding any dominated alternative. Without further knowledge about the decision maker's risk attitude, the 18 alternatives would need to be evaluated combinedly. Furthermore, if risk aversion (hence: second-degree stochastic dominance) can be assumed, the non-dominated set is reduced to twelve alternatives (all except $A_{fr1..2\%,}$ $A_{a0.5\%}$, $A_{a1\%}$, $A_{2\%}$). Risk aversion implies that a decision maker can prefer a less risky to a more risky alternative, even if the expected multi-criteria value is higher for the more risky prospect (Eisenführ et al., 2010). It is a commonly encountered risk behavior (Ananda and Herath, 2005; Pennings and Garcia, 2009), but needs to be validated during preference elicitation.

The top-ranking four alternatives ($A_{f2\%}$> $A_{f1.5\%}$> A_{f2+}~ $A_{f1\%}$) are characterized by medium to high replacement by condition which is favorable regarding the objectives, and especially reliability and intergenerational equity. Costs decrease while reliability increases due to lower failure rates, hence requiring less repairs. The higher replacement rates improve intergenerational equity. The reasoning is similar for A_{cyc80}, but its performance might drop if the average time to failure was much shorter (implying higher failure rates), e.g. due to different material composition or less favorable environmental conditions.

Local sensitivity analysis showed that changes of the weights lead to rank reversals in the non-dominated alternatives and that these are most significant for costs. The value function form had little impact under all scenarios unless all value functions are strongly concave (Figures 16, 17). If extreme preferences such as costs being assigned all the weight or intergenerational equity having zero weight are excluded, the relative ranking of alternatives is rather stable.

The differences in attribute predictions and MCDA rankings under different future scenarios reveal the importance of scenario analysis for strategic rehabilitation planning to inform decision makers about the long-term robustness of different strategies.

For short- and mid-term (i.e. tactical and operational) asset management, these strategies can be extended to account for savings potentially achieved from (1) collaborative asset management with other network infrastructures (e.g. wastewater, gas, telecommunications, road works), and (2) flexible adaptation of annual replacement rates to short-term rehabilitation demands.

5.4.5 Outcome of the case study

For our case study the main results are: If the decision maker is risk-averse (to satisfy the assumption of second-degree stochastic dominance) and unless low costs are most important (very high w_2), $A_{f2\%}$ or $A_{f1.5\%}$ (1.5-2 % annual replacement of oldest pipes in worst condition) is the preferred strategy. If the weights are substantially uncertain, a lower annual replacement rate of 1 % or replacement after the second failure (A_{f2+}) could also be considered, since $A_{f1\%}$ and A_{f2+} are third or fourth-ranked under most assumptions and more robust to weight changes than $A_{f2\%}$ and $A_{f1.5\%}$. Annual replacement of about 1.5 % is typical for larger utilities in Switzerland. Contrarily, the most frequent strategy of small Swiss water utilities and according to (Selvakumar and Tafuri, 2012) also in the USA, namely reactive rehabilitation (A_{ref}), performs well if the only objective pursued is cost minimization. Otherwise, the performance of purely reactive rehabilitation strategies is rather poor and should thus be discouraged. This conclusion is drawn without eliciting weights and risk attitudes, which should be done before deriving final recommendations.

Finally, the decision maker should be cautioned against uncertainty arising from the long-term nature of the predictions (> 40 years) and the limited data basis. The aim should be to embed the strategic rehabilitation plan into an adaptive framework which allows for adjustment of framework conditions, model parameters, and a revision of preferences.

5.5 Conclusions

We suggest a novel approach of combining methods from strategic asset management, failure modeling, decision analysis, and scenario analysis to identify robust long-term rehabilitation strategies for water utilities. The specific problem of pipe failure prediction in small networks with few failure data was successfully overcome by Bayesian estimation of failure model parameters from local data (here: 61 km and 40 recorded failures) and a prior distribution inferred from three larger utilities. The failure modeling procedure extends existing approaches to situations with very limited data, but comes along with important simplifications in data preparation routines and failure modeling which might not be desirable in cases where the available data supports more advanced analyses (sections 5.4.1-5.4.2)

MCDA served as a robust, feasible, and transparent approach to support rational decision making. This is missing in most of the existing approaches, but at the same time demanded by the strategic asset management community (see section 5.1.4). In this paper, we hope to have demonstrated the usefulness of integrating systemic approaches borrowed from decision analysis into engineering modeling approaches. Moreover, we found the combination of MCDA with scenario planning to be highly beneficial. Scenario planning is a new trend in the decision sciences (Montibeller et al., 2006; Stewart et al., 2013). It allows to consider the often neglected future uncertainty regarding the alternative outcomes, as well as assessing the robustness of the alternative rankings under different preferences. Local sensitivity analysis over diverging preference assumptions showed that, in this case, the alternative ranking is most sensitive to the stakeholder's weighting of the objectives, especially under the Boom scenario. Our approach can be easily adapted to other objectives and / or attributes so that alternatives are compared based on aspects that matter to the respective decision maker(s).

Although purely reactive repair (A_{ref}) is the cheapest alternative in terms of rehabilitation costs, it can be expected to perform less well in cases where damage costs to tertiary parties are included. Because its performance regarding intergenerational equity and system reliability is additionally poor, following a proactive rehabilitation alternative is preferable to the still (too) common reactive rehabilitation practice of water utilities.

5.6 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 Adrian Rieder, Sébastien Apothéloz, Christoph Meyer, Thomas Weyermann, and the stakeholders from the case study area for their cooperation and contributions. We also thank Fichtner IT Consulting for giving access to the FAST software and Markus Schmies and Holger Pietsch at Vesta Business Simulations for their support during the implementation of the rehabilitation model. Discussions with Daniela Fuchs-Hanusch and Markus Günther from TU Graz led to substantial improvements of the data preparation routine. Finally, we thank three anonymous reviewers and our colleagues at Eawag, particularly Nele Schuwirth and Carlo Albert, for their valuable inputs.

5.7 Appendices

Appendix A) Length homogenization procedure

Since GIS data was not provided, pipes were left as is, merged or split as follows:

Leave: Pipes and their recorded failures are left unchanged if the pipe length is between 100 and 200 m.

Split: Pipes longer than 200 m are split into separate pipes of equal length and their failures randomly assigned to a position on the pipe. The position of the first failure is sampled from a uniform distribution over the length of the pipe before splitting, while subsequent failures are sampled from a normal distribution N(μ= 0, σ= 75) around the position of the first failure, implying that roughly 95 % of the failures fall within 150 m of the previous. Sample points leading to positions outside the extensions of the pipe before splitting are rejected.

Merge: Pipes shorter than 100 m are merged by subsequently adding pipes of equal laying date, material and diameter subsequently until a further addition would lead to exceed a total of 200 m. Merged pipes are thus not necessarily neighboring pipes. Pipe failures are added from the merged pipes and failure orders recalculated according to their order of occurrence after reassignment. Failures on the same date on one pipe are deleted.

Appendix B) Future scenarios

Future network expansion is linked to population increase. Based on the scenario numbers defined in a stakeholder workshop for the case study region, including water supplier D[12], population increase was assumed as:

$$Population\,[inh.]: P = P_0 \cdot e^{(T-T_0)\cdot cr} \qquad \text{(A.1)}$$

P_0 is the population in the reference year T_0 (here: P_0= 9'540 inhabitants in T_0= 2010), T the evaluation year (e.g. 2050), and cr the scenario-dependent population change rate. Future network expansion after 2010 is derived thereof, assuming a current ($l_{P,0}$) and future per person expansion length l_P, and two adjustment factors g_1 and g_2 to account for changing diameter proportions in the overall pipe network:

$$Expansion\,[m]: E = g_2\left(l_P \cdot P_0 \cdot e^{(T-T_0)cr\cdot g_1} - l_{P,0} \cdot P_0\right) \qquad \text{(A.2)}$$

Network expansion is assumed as PE and DI2 only, being the most strongly increasing materials during recent years in Switzerland[13]. Diameters ≤ 150 mm are assumed to expand as PE pipes, larger diameters as DI2 pipes. The detailed parameters of the four future scenarios are stated in Table 23.

---

[12] Lienert, J., Scholten, L., Egger, C., Maurer, M., 2013. Structured decision making for sustainable water infrastructure planning under four future scenarios. Under review.

[13] SVGW, 2006. Statistische Erhebungen der Wasserversorgungen in der Schweiz, Zürich, Schweizer Verein des Gas- und Wasserfaches.

Table 23: Main characteristics of the four future scenarios*

Name	Socio-economic situation	Population and network expansion			
		c	**l_P [m/inh.], $l_{P,0}$[m/inh.]**	**g_1**	**g_2**
Status Quo	As today: rural region near Zurich with extensive agriculture, leisure areas and nature protection zones. Real income change: +0.4 %/year	No change	No change	No change	No change
Boom	High prosperity, dense urban development, strong nature protection, new transportation. Real income change: +4.0 %/year	$5.284 \cdot 10^{-2}$	l_P: 3.641, $l_{P,0}$: 9.513 Higher building densities lead to less pipes per capita	**<DN150:** 0.5447 **DN150-250:** 0.8643 **> DN250:** 0.6698	1
Quality of life	Prosperous region with moderate population growth, limited expansion of building areas, high environmental awareness. Real income change: +2.0 %/year	$4.558 \cdot 10^{-4}$	$l_P = {}_{P,0}=9.513$ Similar building densities as today.	1	**< DN150:** 0.64 **DN150-250:** 0.32 **> DN250:** 0.04
Doom	Economic recession causes strong financial pressure on municipal budgets, slight population decline but no system expansion / deconstruction. Real income change: -1.5 %/year	$-1.282 \cdot 10^{-3}$	No change	No change	No change

* The mean income in 2008 was 64'575 CHF. With 0.4 % observed increase, the income in 2010 is 65'093 CHF

Appendix C) First-degree stochastic dominance- risk profiles

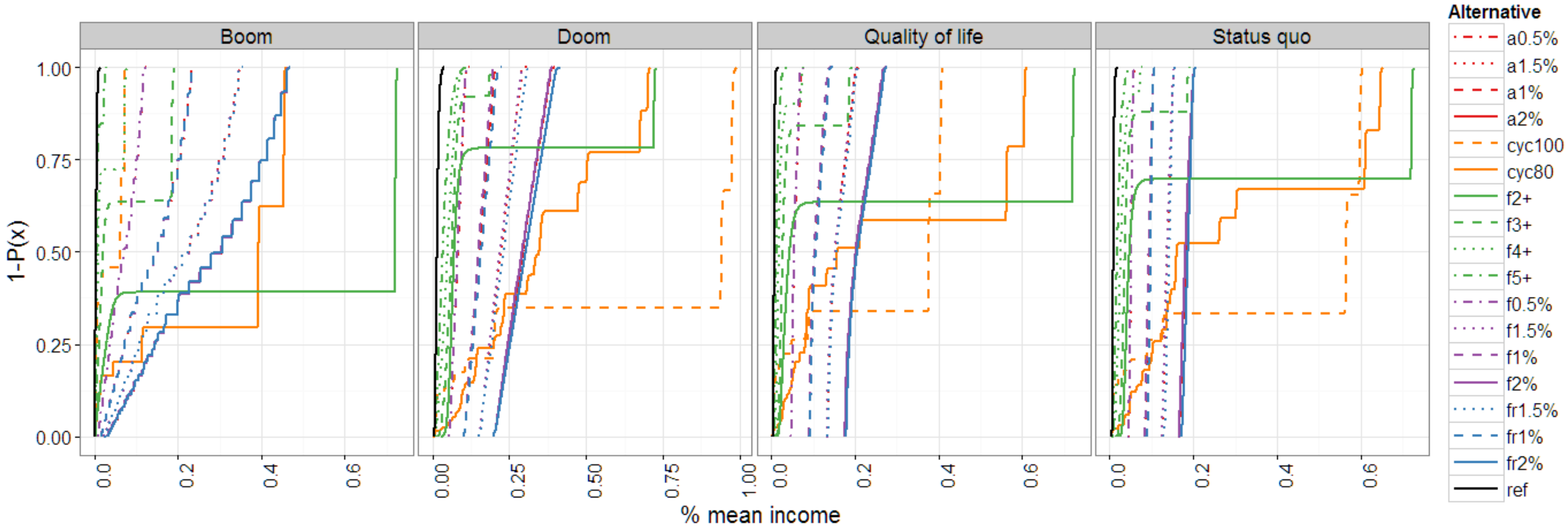


Figure 18: Risk profiles of the alternatives for costs (attribute: % of the mean annual income) over the time horizon 2010-2050

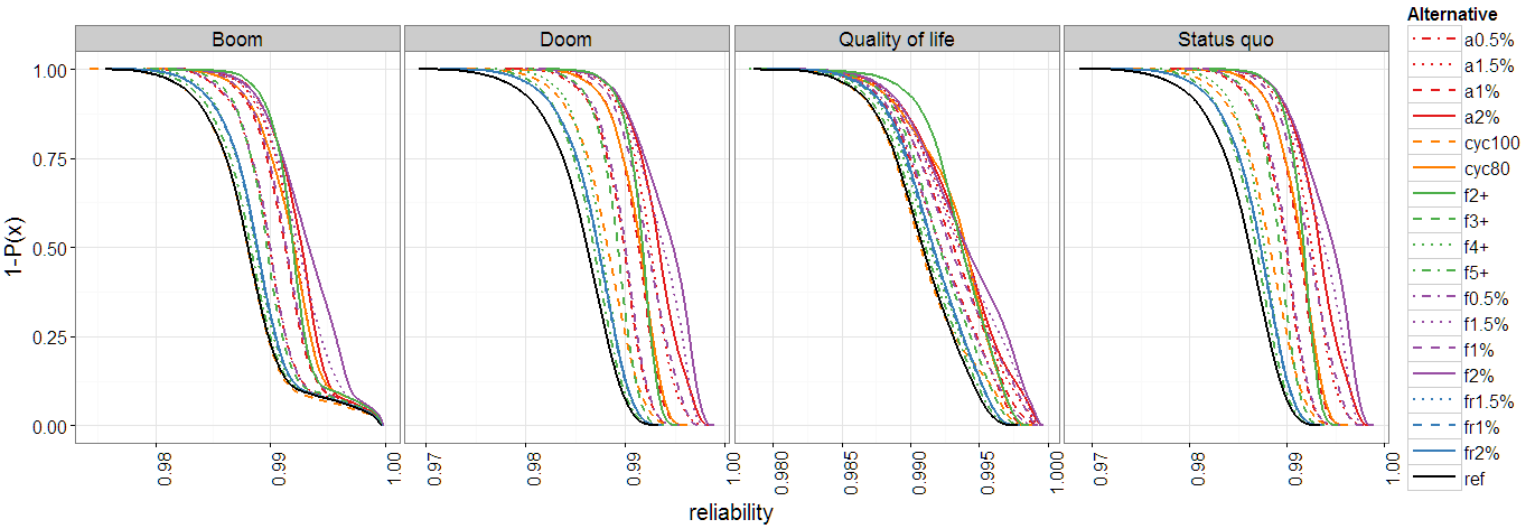


Figure 19: Risk profiles of the alternatives for reliability (attribute: system reliability) over the time horizon 2010-2050

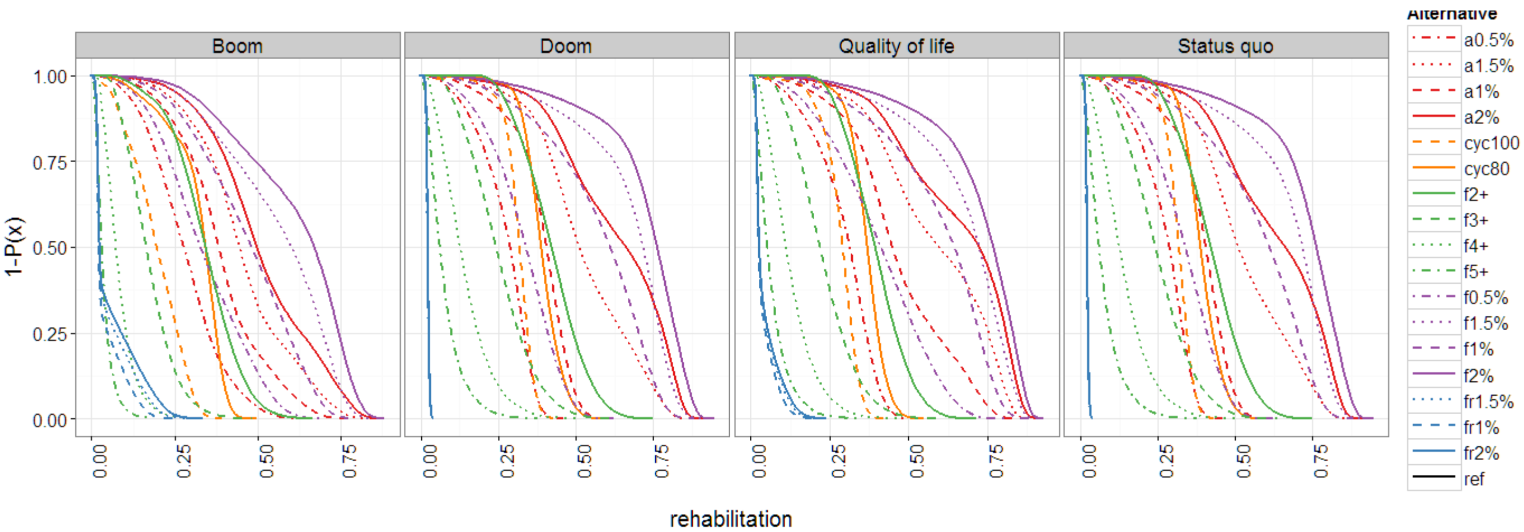


Figure 20: Risk profiles of the alternatives for intergenerational equity (attribute: degree of rehabilitation in %) over the time horizon 2010-2050. The outcome for $A_{ref\ is}$ equals zero (not shown).

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

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

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

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

Shortened version submitted to European Journal of Operational Research.

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.

6.1 Introduction

6.1.1 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 and Raiffa, 1993), has considerably increased in the environmental sciences (Ananda and Herath, 2009; Huang et al., 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 et al., 2005; Joubert et al., 2003), b) use of linear marginal value functions (Raju and 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 et al., 2004; Martin et al., 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 et al., 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 et al., 1997; Durbach and Stewart, 2011, 2012b; French, 2003; Kangas and Kangas, 2004; Keeney and Raiffa, 1993; Stewart, 1995, 2005), it is commonly not considered in practice.

6.1.2 Sources of uncertainty

Different sources of uncertainty in MAUT 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. (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 et al., 2001; Borcherding et al., 1991; Morton and Fasolo, 2009; Weber and 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., 2012a, 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 and 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 and Stewart, 2002; Morton and 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 and 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 et al., 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 and Wright, 2001; Montibeller et al., 2006; Stewart et al., 2013).

(2) Attribute prediction. The uncertainty about the attribute levels of each decision alternative depends on the assessment process. On the one hand, it 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). On the other hand it can stem from the uncertainty of model predictions such as uncertainty of model input / structure / parameters (see e.g. French, 1995; Refsgaard et al., 2007; Walker et al., 2003).

(3) Hierarchical aggregation function. The multi-attribute value or utility function is typically structured hierarchically (see later example, Fig. 21). 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 et al., 2010; Keeney and 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 and Douglas, 1928), minimum-models, or mixtures of these (e.g. Langhans et al., 2013; Langhans et al., submitted; Schuwirth et al., 2012a).

(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 and Schoemaker, 1985; Wakker and Deneffe, 1996) or obtained from converting value functions to utility functions given a specific intrinsic risk attitude (Dyer and Sarin, 1982). Again, several biases are known. For assigning values: *scope insensitivity* and *reference point effects* (e.g. Morton and Fasolo, 2009), and for the assessment of utilities (Bleichrodt et al., 2001; Cox et al., 2012; Eisenführ et al., 2010): *non-linear weighting of probabilities* (Kahneman and 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 et al., 2007; Cox et al., 2012; Rabin, 2000; Schmidt et al., 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 and Fasolo, 2009; Weber and 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 et al., 2001) and ex post corrections (Jacobi and Hobbs, 2007) have been suggested to minimize the splitting bias.

6.1.3 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 et al., 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 and Bosque Sendra, 2004; Saltelli et al., 2006; Saltelli et al., 1999a). GSA has been suggested to support decision makers in the analysis of results from MCDA studies (Mustajoki et al., 2006; Saltelli et al., 1999a), but only applied in few cases (e.g. solid waste management: Gómez Delgado and 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 et al., 2006; Mustajoki, 2012; Mustajoki et al., 2006; Raju and Pillai, 1999) or a combination of aggregation parameters and attributes (e.g. Gómez Delgado and Bosque Sendra, 2004; Gómez Delgado and 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. (2012a) 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 et al., 1998; Lahdelma and 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.

6.1.4 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 et al., 2013; Ruth et al., 2007; Sharma et al., 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 and Brown, 2009). Additionally, alternative forms of utility governance can be chosen, e.g. regionalization or (partial) privatization to achieve higher efficiency and professionalism (Dominguez et al., 2009; Lieberherr et al., 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).

6.1.5 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. (submitted) . 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 et al., 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 6.2. 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., submitted), and preference elicitation including consistency checks. In section 6.3 we present the elicited preferences, attribute predictions, and resulting rankings of alternatives under uncertainty and for four future scenarios. The results of the global sensitivity analysis for one exemplary stakeholder are shown in section 6.4. The results from 6.3 and 6.4 are discussed in section 6.5 and conclusions are drawn in section 6.6.

6.2 Material and methods

6.2.1 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 D1). 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 (SH9 and SH10), although these were judged less important for local planning processes. Together, they represent different entities and decision-making levels, summarized in Tab. 24:

Table 24: 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 Fig. 21 was developed in individual interviews and a stakeholder workshop (Lienert et al., submitted). 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" , Fig. 21. 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 Tab. D2.1 (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.

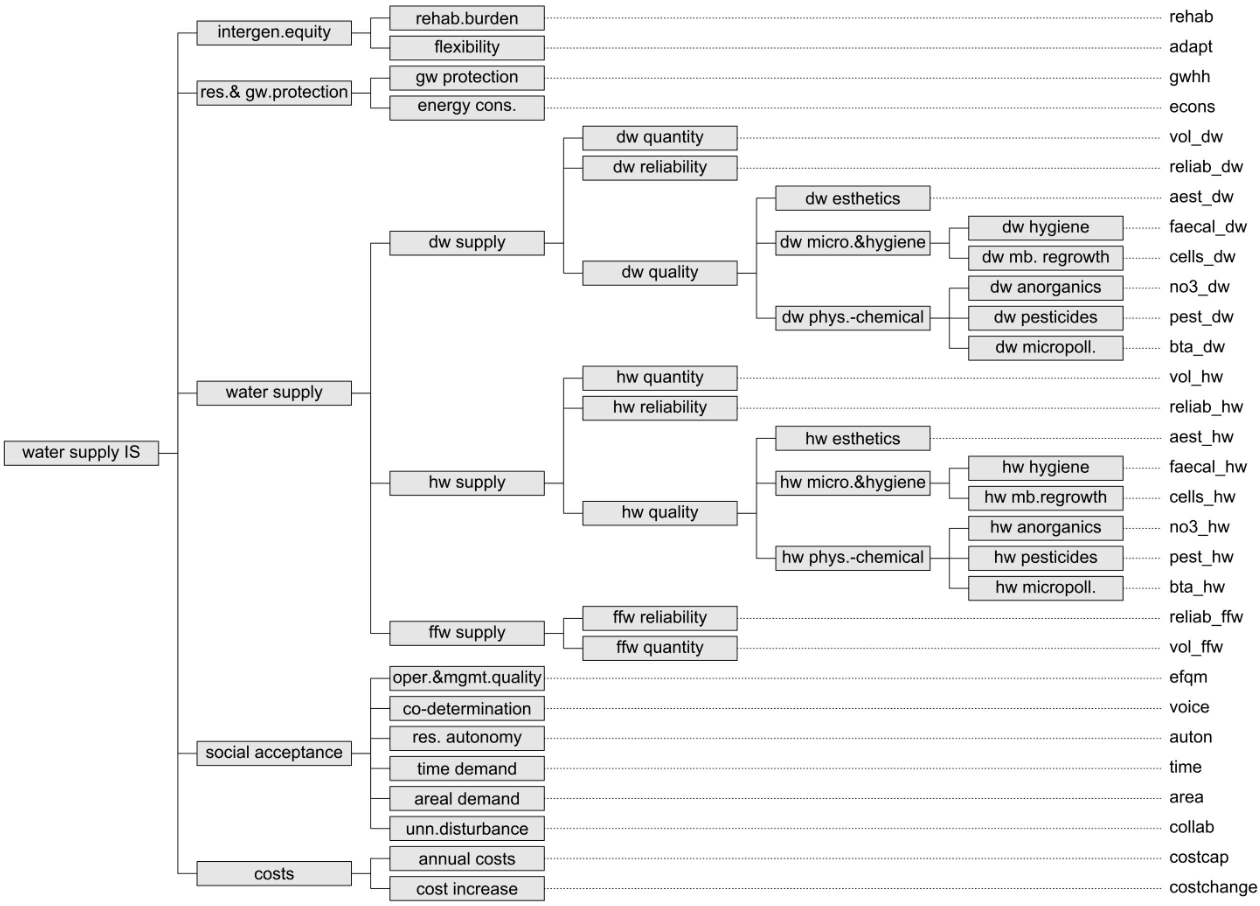


Figure 21: Objectives hierarchy for achieving the overall objective of 'good water supply infrastructure'. Boxes show the fundamental objectives which are connected to the corresponding sub-objectives or attributes (end of the dotted line, right edge of the plot). For more details and the meaning of the abbreviations see Tab. D2.1 (SI).

Decision alternatives

Altogether eleven decision alternatives were generated in a stakeholder workshop, see chapter 2 (Lienert et al., submitted). 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 options, 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 D3.

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., submitted) 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 et al., 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 and Wright, 2001). That means those preferences are used, which current decisions are based on. 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.

6.2.2 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 and 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, Fig. 21) 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.

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., 2012a), 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 Tab. D2.1), 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 Fig. 21). 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., 2012a). 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 and 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., 2012a), 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 inacceptable. The stakeholder specified whether the threshold should affect the overall assessment of the alternative (i.e. would it be inacceptable, 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.

6.2.3 Preference modeling

The stakeholders' preferences are described by individual multi-attribute 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 and Raiffa, 1993; weighted arithmetic mean) and the Cobb-Douglas model for non-compensatory aggregation (Cobb and Douglas, 1928; weighted geometric mean). Their mathematical functions are given in Table 25.

Table 25: Used multi-attribute 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 and Sarin, 1979; Keeney and Raiffa, 1993)
Cobb-Douglas model	$V(v) = \prod_{i=1}^{m} v^{w_i}$	(Cobb and 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 and Raiffa, 1993):

$$U(V) = \frac{1 - e^{-rV}}{1 - e^{-r}} \quad (39)$$

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

$$v_j(x_j) = \begin{cases} \frac{1 - e^{-c_j \tilde{x}_j}}{1 - e^{-c_j}}, & c_j \neq 0 \\ \tilde{x}_j, & c_j = 0 \end{cases} \quad (40)$$

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

6.2.4 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)

$$EU(a) = \int p_a(x) \cdot u(x)\, dx \quad (1)$$

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 and 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 Tab. D2.1 and Tab. D3.2 for details and distributional assumptions. In the second case, the experts' estimates (intervals) were interpreted as 90 % confidence intervals of a normal distribution, the lower range as 5 % quantile, the upper as 95 %. From these, we obtained the mean and standard deviation. 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 could be appropriate at each aggregation node, if preferential independence of objectives holds. Hence, the aggregation function is

$$V(a_k) = \alpha_k \cdot V_{add} + (1 - \alpha_k) \cdot V_{cd} \tag{2}$$

and the mixture parameter a_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. D4.12a–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 ~ *Unif[-10,10]*

Multi-attribute utility function

Since we did not elicit the aggregation parameters for utilities but only the parameters to aggregate values (section 6.2.3), 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 ~ *Unif[-10,10] (Eq.39).*

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 ± 1.96 times the standard deviation σ_i to cover the 95 % interval (SI, Tab. D4.1). 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.

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, Tab. D4.6). 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,

Fig. D4.12e). SH10 specified a step function for the attribute "% utilization of groundwater recharge" with absolute certainty (SI, Fig. D4.12b), which we used instead of an exponential function.

6.2.5 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; Saltelli et al., 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):

$$S_z = \frac{Var_{\theta_z}(E_{\boldsymbol{\Theta}_{-z}}(Y|\theta_z))}{Var(Y)} \tag{1}$$

$Var(Y)$ 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,

$$S_{Tz} = \frac{E_{\boldsymbol{\Theta}_{-z}}\left(Var_{\theta_z}(Y|\boldsymbol{\Theta}_{-z})\right)}{Var(Y)} = 1 - \frac{Var_{\boldsymbol{\Theta}_{-z}}\left(E_{\theta_z}(Y|\boldsymbol{\Theta}_{-z})\right)}{Var(Y)} \tag{2}$$

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 6.2.3), 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 (Fig. D6.1). 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 "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 (Tab. 26).

Table 26: 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 focussed 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 tos?

6.2.6 Implementation

Most of the preference and uncertainty modeling was implemented in R (R Development Core Team, 2011). The R package *utility* (Reichert et al., 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 and Varadhan, 2011), *DEoptim* (Mullen et al., 2011), *adaptMCMC* (Scheidegger, 2012), *sensitivity* (Pujol et al., 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).

6.3 Results of the case study

6.3.1 Attribute outcomes

Attribute outcomes in the Boom scenario differ substantially from those in the other three scenarios (Fig. D3.1a-f). This is the case for attributes whose performance is strongly linked to the scenario assumptions (section 6.2.1, "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 6.3.3 where appropriate.

6.3.2 Stakeholder preferences

Weights

The top-level objective ***"good water supply"*** (Tab. D4.1, and Fig. D4.1) 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 (Fig. D4.2), "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 (Fig. D4.11).

The ranks and weights of the remaining top-level objectives were more divergent. ***"High resources and groundwater protection"*** (Fig. D4.1) was rated highest by SH10 (w= 0.4–0.48, first) and lowest by SH9 (w= 0.12–0.18, fifth). Its sub-objective "natural groundwater regime" was considerably more important than "low energy demand" for all stakeholders except SH5 (Fig. D4.8), i.e. the performance of alternatives for this objective is driven by the utilization of groundwater resources. Similarly, the weight for ***"high intergenerational equity"*** (Fig. D4.1) is subject to high variation (w= 0–0.38). In the extreme cases, it was either irrelevant (w=0, SH10) or second-ranked (SH4, SH5, SH6). Seven stakeholders (all except SH3, SH7), considered its sub-objective "high realization of the rehabilitation demand" more important than "high flexibility" (Fig. D4.9). Finally, the objective ***"low costs"*** (Fig. D.4.1) ranked either third or fourth- (w= 0.07–0.29) for nine stakeholders, and second for SH1 (equal to "high resources and groundwater protection"). The weights of its two sub-objectives "low annual costs" and "low cost increase" were also very different (Fig. D4.10). Due to the weight variations of these three top-level 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. Interestingly, the ranking of objectives in the online surveys was very similar compared to the face-to-face interviews, but the judgment about the relevance of objectives was not (Tab. D4.3). 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 information about 172 value functions, but the shape was only elicited in detail for 21 of them (Figs. D4.12a–f; summary Tab. D4.4). 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), see Tab. D4.5. 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 only 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 (Tab. D4.6). They concern either specific attributes, or a perceived loss / deterioration compared to the current situation. They most commonly addressed drinking water quality concerns, specifically "days per year with hygienic concerns" (mentioned by 8 of 10 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 Tab. D2.1). This was considered by increasing the stated interval of the certainty equivalents, so that all stakeholder preferences could be reconciled with the assumption of utility independence.

6.3.3 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) (Fig. 22; Tab. D5.2). 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, see Tab. 28 (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 classified into two groups by their rank distribution patterns. They are marked either by a circle or a triangle (Fig. 22).

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 (Fig. 22). 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" (2, section 6.2.4, "Acceptance thresholds and individual adjustment"; and Fig. D4.5; Tab. D4.6). According to the attribute predictions, those of A9 could sometimes exceed threshold of 0–10 d/a, thus explaining its low performance (Fig D3.1b). 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 %, Fig. D3.1a), "flexibility of technical extension or deconstruction of infrastructure" (*adapt*, ca.45–65 %, Fig. D3.1a), "system reliability" of drinking, household, and firefighting water (*reliab_dw/ hw/ ffw*, all <0.01, Fig. D3.1b/c/e), "d/a with esthetic impairment" of drinking water (*aes_dw*, 0–10 d/a, Fig. D3.1b), and the cost attributes (*costcap*, *costchange*, ca. 0.01 % of mean income and < 0.8 % increase per year, Fig. D3.1f). Its poor performance with respect to "utilization of groundwater recharge" (*gwh*h, ca. 80–150 %, Fig. D3.1a) in the Boom scenario explains why it is not one of the best alternatives in that case. In the Boom scenario, A1b, and A2 performs similarly well regarding most of these objectives, and outperforms A6 regarding "utilization of groundwater recharge" (*gwhh*, <15 %), and some of the household water quality attributes *(aes_hw, faecal_hw, cells_hw*, Fig. D3.1b-d).

A2, A6 and A9, are technically very similar (see Tab. D3.1). All have a centralized water supply, but the dimensioning of new pipes is done according to peak demands from households alone (not considering peak demands for firefighting). As a consequence, pipe diameters can be reduced which has a positive effect on costs and water quality. Firefighting water is additionally provided by decentralized firefighting water tanks. A6, however, foresees decentralized rainwater storage and use for purposes that do not require potable water (e.g. toilet flushing), causing higher esthetic and hygienic impairment of household water. The high "utilization of groundwater recharge" is explained by the seasonality of available rainfall in connection to a limitation of water imports from external sources to 10 % of the total water demand. Consequently, groundwater resources are extensively used in A6 to satisfy the water demand of the eight-fold larger population in the Boom scenario when rainwater sources are depleted. As "resources autonomy" (*auton*) was only judged of little importance (low weights, if not discarded, see section 6.3.2, Fig. D4.11), it might make sense to trade the currently high autonomy against larger external water imports and thus improve the performance of A6 in the Boom scenario. This does not explain the comparatively bad performance of A9, however, which can be linked to its minimal rehabilitation, operation, and maintenance regime. This practically "do nothing"- strategy leads to a deterioration of water quality (*aes_dw/ hw, faecal_dw/ hw, cells_dw/ hw*), but also non-realization of the rehabilitation demand (*rehab*), and less reliability of drinking and household water supply (*reliab_dw/ hw*), see Fig. D3.1a–e.

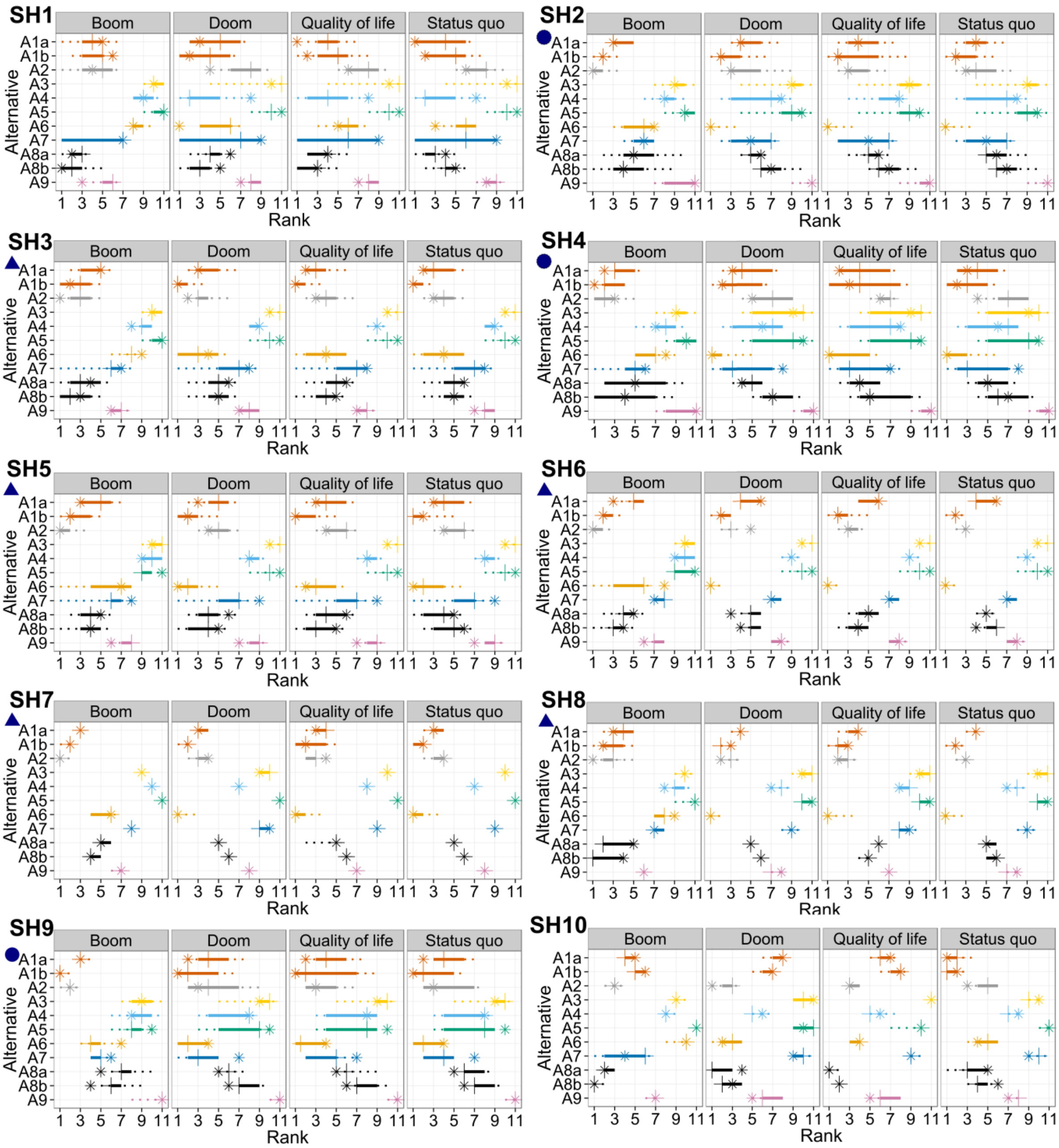


Figure 22: Ranking of eleven alternatives for water supply infrastructure for all ten stakeholders and four future scenarios considering uncertainty. "1" indicates the best rank and "11" the worst. Thick, solid lines with vertical dash represent the 25 %, 75 %, and 50 % quantiles (lower quartile, upper quartile, median), dotted lines the 5 % and 95 % quantiles. Stars indicate the expected ranking with the 'usual' simplification assumption (section 6.2.4). 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 (Fig. 22). 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 is not easily explained by one single preference characteristic. None of the stakeholders stated strict acceptance thresholds (Tab. D4.6), which might influence the ranking.

The potentially best and worst alternatives for the triangle group are technically very different. Alternative A1b, and A8b are conventional centralized water supply systems, but with higher spatial integration and more advanced treatment than in today's system in the case study area. New residential pipes in A8b are dimensioned on reduced water flows ("self-cleaning networks"; Vreeburg et al., 2009) which are also cheaper than conventional networks (see *costcap* and *costchange*, Fig. D3.1f). In addition, moderate (A8b) to extensive (A1b) efforts are spent on rehabilitation, operation, and maintenance of the system, having positive effects on water quality (*aes_dw/ hw, faecal_dw/ hw, cells_dw/ hw*), intergenerational equity (*rehab*), and the reliability of water supply (*reliab_dw/ hw/ ffw*), but lead to high costs (*costcap, costchange*) see Fig. D3.1a–f. Opposed to that, the poorly performing alternatives A3, A4 ("Decaying centralized infrastructure, decentralized outskirts", light blue), and A5 are decentralized water systems with high spatial fragmentation, besides minimal rehabilitation, operation, and maintenance efforts. The effects of this rehabilitation, operation and maintenance regime imply the same negative impacts as described above for A9 (deterioration of water quality, non-realization of the rehabilitation demand, lower system reliability). Additionally, the esthetic quality of household water might suffer from impairments (*aes_hw*, Fig. D3.1c) due to less treatment / and or longer residence times.

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 (Fig. D3.1b–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 Fig. 22). 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 (Fig. D4.12f). 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" (Tab. D4.1). 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 (Fig. D4.12b), otherwise the whole alternative is inacceptable (see ATs, Tab. D4.16). This explains the poor ranking of A6 in the Boom scenario (likely exceeding 100 %; Fig. D3.1a). 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; Fig. D3.1b-c, Fig. D3.1f).

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 Fig. 22. 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 Tab. D5.3). 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, see Tab. 28 (Appendix).

6.4 Results of the global sensitivity analysis

The first (S_z) and total order (S_{Tz}) sensitivity coefficients for the five analysis layouts (cf. Tab. 27) 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 (Tab. D6.1, Fig. D.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 27 ($\sum\theta_z$). 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 Fig. 24), 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 (Tab. 27; see "∑Rank 1–5"). All other parameters are much less relevant, with main effects < 2 % and considerably lower interaction effects (smaller red bars in Fig. 24). 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 (Tab. 27, lower part).

The high importance of the utility curvature parameter *r* can be explained by the distributions of the values, see Fig. 23. 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.

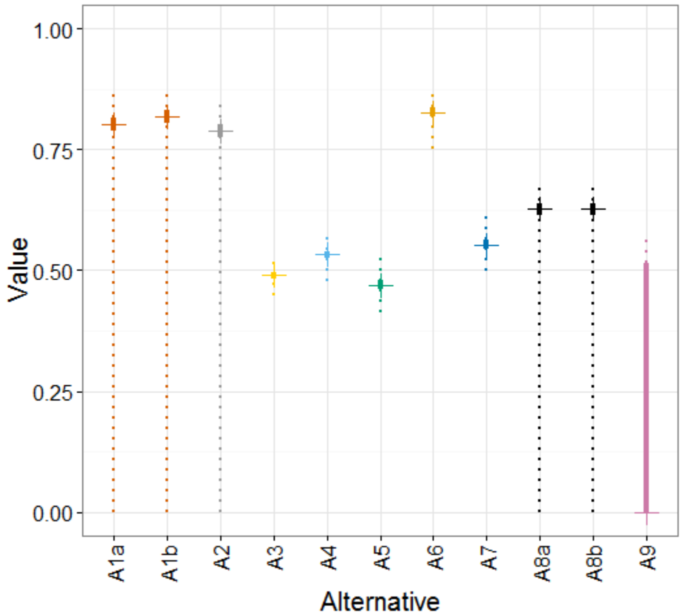


Figure 23: Overall value of the alternatives using the mean preference parameters for SH2 in the Status quo scenario. The 50 % quantile is represented by horizontal bars, the upper and lower quartile by the solid vertical line and the extremes by the dotted vertical lines.

If the ATs are not accounted for, uncertainties of other parameters become more influential, reflected in the respective sensitivity coefficients ("SH2_SQ_noAT", Tab. 27). 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 the risk attitude is much lower (negligible main effect, low interactive effect; see also Fig. 24 "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 Fig. 24, 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 27: Sensitivity coefficients of five analysis layouts. The assumptions underlying the five layouts are summarized in Tab. 26. 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. *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 Fig. 21). 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 Tab. D4.1)

	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
	$\sum\vartheta_z$	**0.771**		**$\sum\vartheta_z$**	**0.524**		**$\sum\vartheta_z$**	**0.386**		**$\sum\vartheta_z$**	**0.483**		**$\sum\vartheta_z$**	**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	

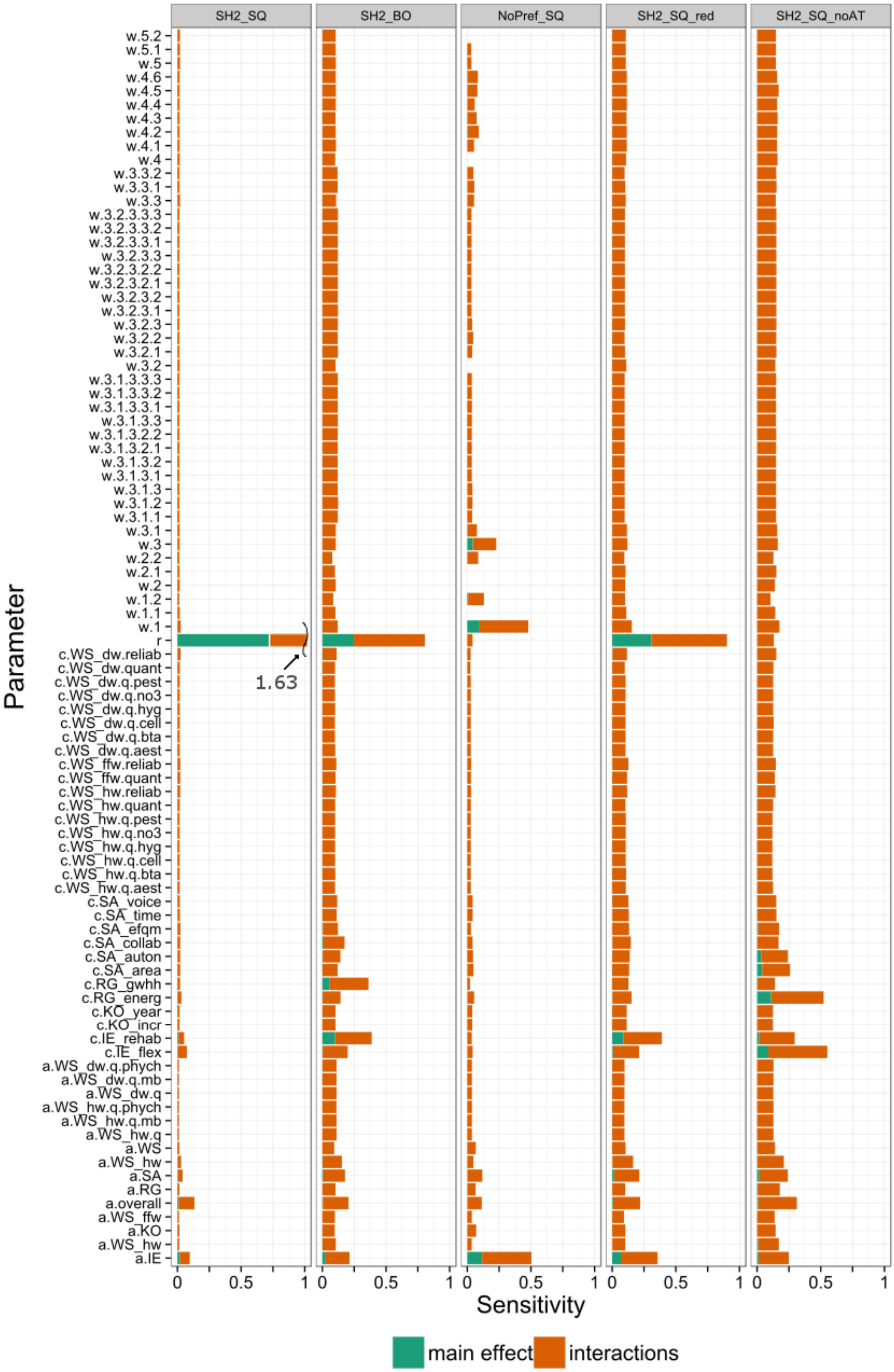


Figure 24: First (S_z) and total order (S_{Tz}) sensitivity coefficients for the analytic layouts presented in Table 26. The S_z (“main effect”) are represented by green bars, S_{Tz} (“interactions”) by red bars. Description s. Tab. 27.

6.5 Discussion

6.5.1 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. Contrary to the often assumed high importance of costs in practice, costs were only the third or fourth ranked-objective for eight out of ten stakeholders. This is in line with findings from other MCDA studies in the water field, where costs were relatively less important than wastewater quality in the removal of pharmaceuticals from hospital wastewater (Lienert et al., 2011) or the quality and reliability of irrigation water supplies (Hyde et al., 2011).

No single best or worst alternative for all stakeholders or scenarios could be identified, as the individual ranking distributions differed and sometimes overlapped (section 6.3.3). 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", Tab. D3.1), 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. This is interesting, as the potential reduction of pipe dimensions in residential areas is often rejected by the perceived need to supply large amounts of firefighting water (enforced by building codes that specify the minimum standards for constructed objects) to withstand any worst-case scenario. Large, decentralized firefighting tanks were foreseen in A2, A6, and A9 to meet this requirement, although the provided flows from smaller-diameter pipes are usually sufficient for reliable fire-fighting in residential areas, e.g. as reported from the Netherlands (Vreeburg et al., 2009). 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. Although additional rainwater use in A6 reduces water demand for non-potable purposes, the demanded amounts of water are very high in the Boom scenario (eight-fold the population of 2010) and lead to over-exploitation of the local groundwater resource. The stakeholder preferences (section 6.3.2) 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 under the Boom scenario.

The fully decentralized alternatives A3 ("Fully decentralized") and A5 ("Decaying infrastructure everywhere") had poor ranks given our assumptions. They are characterized by in-house water treatment in combination with highly fragmented management and minimal rehabilitation and maintenance efforts. They had particularly poor outcomes regarding the realization of the rehabilitation demand and reliability of drinking and household water supply which furthermore leads to a deterioration of water quality. The performance of the decentralized alternatives given more favorable management regimes needs to be analyzed to gain further insights into the importance and effect of rehabilitation and management on overall alternative performance, as implied by the better ranking of A7 ("Mixed responsibility, fully decentralized with onsite treatment", moderate efforts for operation, maintenance, and rehabilitation). An analysis of the aggregated values of the alternatives instead of the rankings based on expected utilities alone will provide more insights into how much less preferred alternatives A3 and A5 really are and if they can be upfront discarded based on this information.

6.5.2 Preference elicitation

We find the presented two-step procedure consisting of an online survey and a face-to-face interview (section 6.2.2) 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 and Fasolo, 2009; von Nietzsch and Weber, 1991 (in Eisenführ et al., 2010); Weber and 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.

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., 2012a), 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 6.3.1, "Preference thresholds and individual adjustments", and 6.3.2). 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 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).

6.5.3 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 6.2.4). 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 and 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 6.2.4). 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.

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).

6.5.4 Global sensitivity analysis

We also showed how GSA can be used to explore which of the uncertain parameters have the largest impact on the results (section 6.4). 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., 2012a; Keeney and 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 and Herath, 2009; Hajkowicz and Collins, 2007; Mendoza and 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 6.2.2) 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 and 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).

6.6 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.

6.7 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, 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.

6.8 Appendix

Table 28: 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 6.2.4), 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 Tab. D.11-12 (SI)

	USUAL SIMPLIFYING ASSUMPTIONS: Individual ranking										Mean of SH		UNCERTAIN PREFERENCES: 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

7 Conclusions and outlook

7.1 Conclusions

The aim of this thesis was to develop a methodology for more comprehensive and integrative long-term water supply infrastructure planning under uncertainty. It should explicitly consider different sources of uncertainty and be tested it in a real case study in Switzerland. Out of four challenges in current water infrastructure planning, three are addressed by the approach presented in this work:

The lacking knowledge about water infrastructure condition and rehabilitation demand is addressed by pipe survival and failure models which are adapted to the data situation, as presented in chapter 3 and 4. The rehabilitation demand is determined using these models and combining them with rehabilitation models to assess the performance regarding specific objectives and support strategic asset management (SAM). To evaluate discrete rehabilitation strategies based on a decision maker's preferences, SAM is then embedded into a multi-criteria decision analysis (MCDA) framework using multi-attribute utility theory (MAUT), see chapter 5. If applied in a multi-stakeholder context, the use of a facilitated MCDA approach for water infrastructure planning is able to overcome both the lack of multi-stakeholder involvement, and transparency in current decision making (chapter 6). While uncertainties arising from uncertain inputs and model choice are explicitly considered by means of probabilistic modeling, future uncertainties can be explored by the integration of the above methods with scenario planning. The number and type of alternatives can be adjusted at will, and the system of objectives be extended or reduced to fit the specific decision problem. It has been shown in chapter 2 that it is not only possible, but also that stakeholders are keen to overcome the narrow-minded extrapolation of the status quo, consider wider objectives and unconventional alternatives besides future dynamics in their decision making. The presented approach provides the means to do so, but needs further research and adaptation to become feasible given the available resources (see 7.2, Outlook).

This leads me to the challenge that is not explicitly addressed by this approach, but rather the symptoms of which are being dealt with: the limited institutional, financial, and professional capacities to perform long-term water infrastructure planning. It will be the task of future science-policy-practice collaboration to decide if the presented approach shall be followed (or parts of it), to define how to make it more practicable, and to eventually accompany its integration into water infrastructure planning.

7.1.1 Pipe survival, failure, and rehabilitation modeling

To improve the current knowledge about the condition and rehabilitation needs, and to predict future performance of the water distribution network, a service life model (predicts when a pipe reaches is end of life) and a failure model (predicts when failures occur) for water pipes were developed, see chapter 3 and 4. Their ability to deal with left-truncated and right-censored data, and also the absence of data due to past replacement is an important step towards less biased predictions of pipe failure rates and service life (e.g. Le Gat, 2009; Renaud et al., 2012; Scheidegger and Maurer, 2012). I demonstrated how Bayesian inference of model parameters can be used for model calibration in situations where Frequentist approaches are unsatisfactory, thus overcoming the problem of small network (sample) sizes.

To obtain suitable priors, a procedure to elicit imprecise quantiles of pipe survival distributions from experts under consideration of common heuristic biases was developed, building on work of O'Hagan et al. (2006) and Rinderknecht et al. (2012), among others. Because the objective was to obtain an (intersubjective) prior of pipe survival for different pipe groups, an aggregation approach is proposed which takes explicit account of individual, but also in-between experts variance. This allows for a better coverage of diverging influence factors of network aging which are not usually considered during data aggregation of survival data. The input uncertainty is described by multi-variate probability distributions of the model parameters which are propagated to the model outputs. The findings add to the growing body of literature in water engineering which considers expert knowledge a valuable source of information by which the absence of data can be overcome or the available knowledge be enriched. The mathematical aggregation of expert estimates allowed to derive individual priors which are consistent with the dominant environmental conditions of the respective water utilities. Furthermore, the intersubjective prior is in agreement with available information on pipe survival (Fuchs, 2001; Roscher et al., 2005; Trujillo Alvarez, 1995). The incorporation of in-between experts variance furthermore permits to acknowledge different environmental and operational framework conditions and proved useful in the context of pipe survival modeling. Failure model priors were obtained using the same aggregation approach, but inferred these from three larger water networks (chapter 5). Furthermore, a novel approach for pipe and failure data homogenization based on propositions by Gangl (2008) and Fuchs-Hanusch et al. (2012) was developed, which is useful in situations where data about the pipe position (e.g. GIS data) are missing. Although it helps to bridge the lacking data situation for strategic management, knowledge about pipe position and location of failures is essential for mid-term (tactic) and short-term (operational) asset management. Lastly, the calibrated pipe failure model was combined with available strategic asset management software for the assessment of performance of selected rehabilitation strategies.

Both the survival and failure model deliver reasonable results and are applicable to small water networks. The priors were shown to be very influential in all examples with small data sets (few pipes, even less replacement failure data). Consequentially, careful formulation of these priors is paramount, because the Bayesian posterior which is used for predictions depends almost exclusively on the prior if local data are scarce. The local conditions can be expected to be better represented as the knowledge (amount of data) increases over time. The prerequisite for this is reliable future data keeping which not only records characteristics and failures of the running system, but also keeps records of decommissioned pipes. The survival model is judged less suitable for making statements about the useful life of pipe groups than the failure model, due to its inability to differentiate between structural and managerial replacement. In contrast, the failure model is built on times between failures which are able to represent the structural aging process, and the number of failures serves as proxy for pipe condition. Its adaptation to the specific data management practice is more demanding though, and can hinder its application since data handling is heterogeneous and consistent data keeping often is not a priority in practice. The Weibull-Exponential failure model used herein is simple and for that reason compelling, but comes along with important simplifications that do not accommodate all particularities of observed pipe failure behavior. An important simplification for example is the assumption of identical distributions of the times between higher order (the second, the third, the fourth…) failures which contradict practical observations (e.g. Kleiner and Rajani, 2001). On the other hand, the inclusion and estimation of additional parameters to better cover these particularities will make model calibration more cumbersome, as the difficulties in identifying parameters for water networks of 220 and 385 km (among the largest networks in Switzerland) indicate, see chapter 5. Theoretically, this is unproblematic; the presented Bayesian approach can overcome this difficulty as shown for smaller utilities. In practice, large enough data sets or experts to derive useful priors for more complex models from may be hard to find. As regards the different pipe groups, the failure model parameters and the intersubjective prior from expert elicitation reveal a consistent durability order: asbestos cement, 2nd generation ductile cast iron (after 1980), 3rd generation grey cast iron (1930-1965), polyethylene, steel, 1st generation ductile cast iron (1965-1980). Pipe grouping in this thesis was either predefined by what characteristics experts were able to differentiate, or by common characteristics of networks, the inferred parameters of which should be aggregated. For individual analyses, other groupings might be more significant and require the definition of priors specific to that grouping.

The rehabilitation model builds on the parameterized failure model to predict pipe deterioration. In contrast to most of the available models (e.g. PARMS Planning: Burn et al., 2003; Aware-P: Cardoso et al., 2012; WilCO: Engelhardt et al., 2003; PiReM: Fuchs-Hanusch et al., 2008; D-WARP: Kleiner and Rajani, 2004; KANEW: Kropp and Baur, 2005; Casses: Renaud et al., 2012), it allows for large flexibility in the exploration of rehabilitation strategies under uncertainty, be it purely condition-based, age-dependent or other managerial strategies. Whereas the uncertainty of the number of future failures is propagated to the results, the uncertainty of the number of failures before the start of failure recording (i.e. the current condition) could not be explicitly propagated due to practical reasons. That would require the possibility to probabilistically specify the current condition which is not implemented. A manual circumvention of this was judged infeasible, given the fact that the network asset data have to be imported one by one into the 72 versions of the model (18 alternatives times 4 scenarios, each running between 2-3 hours), and repeated many times to grant stability of the results. The performed random sampling of previous pipe failures for each individual pipe partly addresses this issue (though indirectly), but does not overcome it.

It should not be forgotten, that the long-term predictions (> 40 years) were derived using comparatively short periods of recorded data (10-15 years). Consequentially, prediction uncertainties, as well as prediction bias or error increase with the length of the prediction horizon. In addition, future dynamics which are known to affect pipe deterioration, e.g. road traffic, stray currents, soil humidity, and duration of frosts, are not considered during modeling. As a consequence, the predictions need regular updating as new data become available and cannot be interpreted in an absolute way.

7.1.2 Multi-criteria decision analysis (MCDA) for water infrastructure planning under uncertainty

Multi-attribute utility theory (MAUT) was combined with scenario planning to overcome the negligence of future dynamics and uncertainty in current water infrastructure planning (chapter 2, 5, and 6). Two studies were performed: (1) the multi-criteria evaluation of rehabilitation alternatives in search of 'good pipe rehabilitation strategies', and (2) the MCDA aiming at achieving a 'good water supply infrastructure' in the "Mönchaltorfer Aa" case study.

The first study (in chapter 5) is of particular interest to the management of aging water supply systems. It demonstrates that strategic asset management can by complemented by MAUT and scenario planning, and in that way serve as a robust and transparent decision support framework for infrastructure managers, as requested by researches in the asset management field (e.g. Alegre, 2010; Giustolisi et al., 2006; Selvakumar and Tafuri, 2012). The propagation of uncertainty from failure modeling to the outcomes of the rehabilitation strategy and evaluation under four future scenarios helps to make more informative long-term decisions. Even without elicitation of stakeholder preferences, the analysis of dominance and sensitivity of 18 alternatives regarding preferences about three objectives ('high system reliability', 'low costs', 'high intergenerational equity') allowed to narrow the set of alternatives down and explore their performance under different preferences (see 7.1.3). The number of objectives can be extended (or reduced) as suitable to support either a single decision maker (e.g. the asset manager) or a group of decision makers (e.g. representatives of the water utility, municipality, and engineering consultancy) in strategic asset management.

The second study (chapters 2 and 6) contributes to past research in both long-term water infrastructure planning, and the application of MAUT under uncertainty. Uncertain future dynamics and uncertainties arising from the elicitation of preference and preference modeling are addressed by linking MAUT, scenario planning, and uncertainty analysis. The combination of MAUT and scenario planning is increasingly common in MCDA (e.g. Goodwin and Wright, 2001; Montibeller et al., 2006; Stewart et al., 2013) and the number of reported uncertainty analyses is also increasing, though usually restricted to the uncertainty of weights and attributes (e.g. Butler et al., 1997; Hyde et al., 2004; Jessop, 2011; Jiménez et al., 2006; Raju and Pillai, 1999; for consideration of other preference parameters see also Durbach and Stewart, 2012; Schuwirth et al., 2012; and Stewart, 1995). This work extends the scope of these approaches by explicitly tackling uncertainty during preference elicitation and modeling and by considering a wider range of sources of uncertainty for the evaluation of eleven alternatives under four scenarios. It allowed important insights into the uncertainty of the resulting rankings, and how much these would deviate from an analysis under 'usual simplification assumptions' (linear single-attribute value functions, additive aggregation, sure weights (best-guess), utilities equal to values). I furthermore demonstrated how an additional global sensitivity analysis (GSA) is useful to determine which of the uncertain preference parameters are most influential regarding their individual or interactive effect on the ranking of alternatives.

The creativity and structuring techniques used to develop the scenario narratives and alternatives during the workshops (chapter 2) evidently encouraged stakeholders to think out of the box, to define regionalized future scenarios relevant to them, and to construct both conventional and unconventional future alternatives. Not only technical alternatives, but also changes in organizational form, spatial extent, and infrastructure management were considered. This is reflected in the extensive objectives hierarchy (44 fundamental sub-objectives), covering (1) water supply-related objectives such as 'high water quality and quantity' or 'high system reliability', (2) sustainability-related objectives e.g. 'high groundwater protection' (environment), 'high intergenerational equity' (society), and 'low annual costs' (economy), and also (3) objectives addressing social acceptance, e.g. 'high quality of operation and management", or 'low unnecessary disturbance from road works'. It can serve as starting point in future water infrastructure planning, e.g. to set up lists of objectives that decision makers can choose from or extend (suggested by Bond et al., 2008 to overcome people's difficulties in generating a comprehensive set of objectives during decision structuring).

A two-step preference elicitation procedure consisting of an online survey and face-to-face interviews was developed and tested with ten selected stakeholders (chapter 6). The online survey was designed to obtain an importance ranking of the objectives, but was also helpful to familiarize the stakeholders with the decision problem. Based on this information, interviews could be focused on the objectives that were rated as most important. During the interviews, imprecise preferences about the scaling factors, value functions, and certainty equivalents were elicited. The additional effort for stating intervals was minimal, sometimes the experts even preferred this (similar to the experience from imprecise expert elicitation in chapter 3). Even though insights about many preference parameters could be gained, no complete preference set was obtained. Consequently, the arising uncertainties of the preference parameters (aggregation function, marginal value function curvature, scaling factors, risk attitude) as well as uncertainties of attribute prediction, were explicitly considered in the preference modeling. Context uncertainties were addressed by performing the uncertainty analysis of alternative rankings under four scenarios. We followed the approach of Goodwin and Wright (2001) when evaluating alternatives in each future scenario. The preferences were assumed as scenario-independent and consequently, only the attribute outcomes depend on the scenario. This stands in contrast to recent works on scenario planning in MCDA (e.g. Karvetski et al., 2009; Ram and Montibeller, 2013; Montibeller et al., 2006) which account for changes of preferences depending on the scenario. Whereas research has shown that preferences may be instable over time (Frederick et al., 2002), elicitation of scenario-dependent preferences seems highly hypothetic given the long time horizons and uncertain future developments. Instead, using current stakeholder preferences in decisions that need to be taken based on the current situation and validating them after a certain time in an adaptive process appears more reasonable. Therefore, the overall robustness of each alternative was derived from changes in the rankings under the four scenarios. The different scenarios were assumed to affect the boundary conditions based on which the attribute outcomes are calculated. These differences are also visible in the resulting rankings of alternatives which could often be directly related to the changes in specific attribute predictions. Based on the available analysis, however, the relative importance of the uncertainty about the attributes as compared to uncertain preferences cannot be determined.

Comparing the obtained rank distributions with the tentative ranking under usual simplification assumptions, a strong divergence between the ranking of alternatives was found for some stakeholders. Nevertheless, the overall recommendations for potential compromise solutions for the case study did not change (see 7.1.3). These results call for greater attention and scrutiny towards the application of usual simplification assumptions in MAUT since they may lead to wrong conclusions (e.g. if the objective is not to identify a compromise solution, but rather understand which alternatives perform best for whom and why). Even though simplifications are necessary for making elicitation feasible, their implications and possible deviations can be verified with the help of uncertainty and sensitivity analysis. The results from GSA for different modeling layouts and using the elicited preferences of one exemplary stakeholder showed that the uncertainty of the model outcome can be strongly reduced if only a few preference parameters were known more precisely (chapter 6). This is an important insight especially in complex MAUT analyses, where not all uncertain parameters can be elicited in detail. GSA helps to identify which uncertain parameters should be elicited in detail to reduce the overall ranking uncertainty, thereby increasing discrimination between alternatives. In addition, a GSA without consideration of any elicited preferences confirmed the weights and aggregation mixing parameters as most influential parameters. This supports the chosen elicitation layout consisting of the detailed elicitation of (interval) weights as well as of selected value functions and independence conditions only.

In this thesis, only the quantitative effects of the ninety uncertain preference parameters were assessed, but not those of the uncertain parameters of the distribution of attribute outcomes (thirty in each future scenario). This would have required considerably more modeling efforts, because much larger parameter samples are needed to obtain stable sensitivity indices (Saltelli et al., 2006; 2010). Therefore, no recommendation can be made regarding how important the reduction of uncertainty in the identified most influential preference parameters is compared to reducing uncertainty in specific attribute parameters. This is an important limitation, because the predicted attribute outcomes are based on very different sources of information and reflect different degrees of uncertainty. For instance, the chemico-phyiscal quality of drinking and household water could not be timely assessed based on the available data and expert estimates. Hence, a uniform distribution over the complete attribute range was assumed, which adds up considerable uncertainty. Although this uncertainty affects all alternatives in the same way (as no distinction between alternatives could be made), it might have an important effect on the overall stability of the ranking of alternatives. The same applies to other attribute outcomes which were roughly estimated with large uncertainties and might well differ between alternatives, e.g. the available volumes of water for firefighting (derived from dimensioning) or the degree of codetermination (assessed by experts). If consequently, the uncertainty of those rough attribute predictions could be considered in al GSA over all uncertain parameters, efforts could be targeted to those aspects which matter and spending considerable resources on extended measurement campaigns or complex predictive modeling be avoided. For future studies the following evaluation procedure in complex MAUT problems is suggested: (1) assessment and prediction of attribute outcomes for each alternative and scenario, (2) elicitation of interval weights and preferential independence conditions (as these limit the choice of possible multi-attribute aggregation functions), (3) uncertainty analysis and ranking of alternatives. If the value and rank distributions strongly overlap and no clear conclusions about which alternatives perform better (or worse) can be derived, an additional (4) GSA is performed to identify which uncertain parameters to assess more precisely (be it preference parameters, as done herein, or extending the analysis to uncertain attribute parameters).

Overall, the combination of MAUT with scenario planning and also of sensitivity and uncertainty analysis was beneficial in many ways. The findings from both studies showed that it does not only provide a transparent, creative, and robust framework for long-term water infrastructure planning under uncertainty, but also makes clear that the importance of considering uncertain future scenarios into today's planning cannot be underrated (e.g. Ferguson et al., 2013; Milly et al., 2008; Ruth et al., 2007; Sharma et al., 2010). Especially the very dynamic socio-economic changes in the Boom scenario showed a strong impact of scenario assumptions on attribute outcomes and alternative rankings. The narrow-minded extrapolation of the current system under stationary assumptions should thus be overcome (e.g. Ashley et al., 2008; Dominguez et al., 2009; Störmer et al., 2009) – not only to ensure higher robustness towards dynamic future developments, but also because the potential for technical system change is highest in very dynamic environments. Likewise, the uncertainty of preference parameters has an impact on the evaluation of alternatives and should thus be considered in preference elicitation and modeling. I see great potential in extending the analyst's toolbox with GSA especially in complex decision problems which involve many uncertain parameters. In addition to its ability to quantitatively assess the influence of individual model parameters, it also measures the importance of interactive effects between parameters. As discussed in Saltelli et al. (2006), LSA is neither able to quantitatively measure the contribution of individual parameters, nor to capture interaction effects. The possible insights from LSA are thus limited and should be interpreted with caution where modeling uncertainties are large and models non-linear.

7.1.3 Findings and recommendations for the Mönchaltorfer Aa case study

Based on the assumptions and modeling for the example utility, the main findings are: Unless extreme preferences apply (e.g. the only objective being cost minimization), the common reactive repair and replacement strategy performs worse than the proactive strategies and should be overcome. The annual replacement of 1.5-2 % of the network by physical pipe condition (poorest first; one of the more common strategies in large Swiss water utilities) were best-ranked for a risk-averse decision maker. If high robustness under different scenarios and uncertain preferences is aspired, 1 % annual replacement by condition or simply replacement after the second failure could also be considered. These findings should be corroborated by elicitation and evaluation of stakeholder preferences and not be generalized until validated in other case studies.

Preference elicitation with stakeholders for the Mönchaltorfer Aa in the second study indicates that the fundamental objective of 'good water supply' is key to 'good water supply infrastructure' and that 'high social acceptance' is the least important given the stated attribute ranges. There is little disagreement between stakeholders regarding these objectives, hence alternatives are most strongly discriminated by objectives linked to good water supply, e.g. good water quality (and especially drinking water hygiene), and high reliability of supply. On the other hand, resources autonomy, time and area demand or other sub-objectives of social acceptance were lowest-ranked or discarded, indicating their little relative importance to the stakeholders as compared to the remaining objectives.

Even though no single alternative performs best or worst for all scenarios, suggestions for potential compromise solutions can be made. A potential compromise is A6 ("Maximal collaboration, centralized"), which performs well regarding many of the highest-weighted objectives (water supply reliability and drinking water quality, technical flexibility, rehabilitation demand, and costs). Technically, it is an adaptation of the current centralized supply system, but dimensioning of new pipes in residential areas is based on household peak demands instead of the (higher) firefighting peak demands, thus reducing their diameter. To fulfill the prevailing firefighting requirements of the cantonal building insurance (GVZ, 2011), decentralized, shared firefighting tanks are allocated. Moderate efforts are spent on management and operation, annually 1 % of the distribution network is replaced by condition. Its spatial organization is more regional than in the current situation, and builds on stronger integration of water utilities (one cooperative running the water services of the four case study municipalities). In addition, rainwater is harvested, screened, and stored to cover part of the water demand for non-potable purposes (e.g. toilet flushing). No more than 10 % of the water is imported from outside the region (e.g. lake Zurich). This is problematic in the "Boom" scenario (highly prosperous economy, booming population growth), because the high water demand of the eightfold increased population leads to an overuse of the local groundwater resources (hence, very poor performance regarding 'high groundwater protection', another very important objective). This can be remediated by allowing more water imports. More water imports imply less 'resources autonomy', but the effect of this on the ranking is probably negligible since the objective was assigned very little importance by the stakeholders. In the other scenarios ("Doom": unfavorable socio-economic development, population decrease; "Quality of life": favorable socio-economy, slight population increase; "Status quo": little socio-economic growth, stable population), A6 has either the first or second overall rank and therefore a high consensus potential.

Interestingly, the alternative with the highest overall rank in the Boom scenario (A2, "Centralized, IKA", see Tab. D.4), is very similar to A6, but does not limit imports from external sources. There is a third alternative, A9 ("Centralized, privatization, minimal maintenance), which is technically very similar, too, but ranks considerably worse under all scenarios. The most obvious reason for this are the envisaged minimal operational, maintenance, and rehabilitation efforts. As a consequence, the outcomes for the 'realization of the rehabilitation demand', 'reliability of drinking and household water supply' are worse and ultimately lead to a deterioration of the water quality.

Minimal efforts for operation, maintenance, and rehabilitation are also characteristic of the two worst-ranking alternatives, A3 ("Fully decentralized") and A5 ("Decaying infrastructure everywhere"). Apart from this, both are fully decentralized (e.g. in-house water treatment, water delivery by lorries), and have highly fragmented management. The responsibility rests mostly on the end users. Their performance regarding the realization of the rehabilitation demand, reliability of drinking and household water supply, and water quality is particularly poor. It remains to be investigated how much the poor ranking of the decentralized alternatives depends on the technology itself or if major part can be explained by the poor rehabilitation and management. The rank differences between A6 and A9 in the centralized case and a fair ranking under all scenarios of A7 ("Mixed responsibility, fully decentralized with onsite treatment", also decentralized, but allocates moderate efforts for operation, maintenance, and rehabilitation; responsibility is shared between utilities and end users) might indicate this. Given its impact on three important objectives (system reliability, water quality, realization of the rehabilitation demand), the importance of a thorough management and rehabilitation of water supply infrastructures cannot be neglected.

Lastly, and to avoid misunderstandings, the large ranking differences between alternatives are not automatically equivalent to large preferential differences. Thus, even though A6 received the highest rank in most cases, the preferential difference between A6 and other alternatives with lower ranks (A2, other centralized alternatives, or maybe also A9) might be small. Before upfront discarding alternatives with presumably mediocre ranking from further discussions with the stakeholders, the aggregate value of the alternatives should be additionally considered, as it provides quantitative information about the preferential differences (but ignores risk attitudes which the ranking based on expected utilities does not). The insights gained from this study can be used to explore meaningful ways of improving potentially best alternatives or to design new promising compromise alternatives for the case study.

7.2 Outlook

Models and modeling, additional analyses

- Due to the high influence of the prior distribution for Bayesian failure modeling in small utilities, more efforts should be spent on cross-validation using data from other networks. Besides having a reliable failure documentation, these networks should be (1) reasonably different regarding their underlying environmental and operational conditions to ensure a good coverage, and (2) large enough to circumvent identifiability issues (based on the experience with the three large utilities in chapter 5: about 100 pipes or roughly 12 km per pipe group and a minimum of 20 recorded failures).

- The failure model should be improved towards a better representation of the time between higher-order failure distributions. This could be achieved, for example, by using separate exponential distributions for higher order failures (e.g. extending the model of Mailhot et al., 2000 with a replacement model that considers the absence of replaced pipes). To better account for different past replacement regimes during model parameter inference, a more flexible selection of the replacement model to consider further aspects such as the age of a pipe at the time of failure (e.g. as done by Le Gat, 2009), or its criticality is desirable.

- The contrary approach to adapting the herein used model to better represent these particularities is to embed existing models which fulfill the requirements into a Bayesian framework so as to allow their application for smaller utilities. The priors would need to be obtained from very large utilities, though. The only currently available model that I know could fulfill these properties (besides accounting for left-truncated, right-censored, and selective survival data) is the model of Le Gat (2009).

- It would be beneficial for the further development of asset management software and / or rehabilitation models to allow for the propagation of the uncertainty about the previous number of failures (hence current network condition), as this can make a difference in the estimated rehabilitation demand and decisions on which assets to replace first, especially in small networks.

- Not only the public water distribution pipes are expected to require increased rehabilitation efforts. Necessary efforts for private household connections might be just as high (Martin, 2009). Therefore, there is high research potential in the analysis of their characteristics and condition, besides the development of failure and rehabilitation models, and their inclusion into strategic asset management programs. The data situation might be even worse than for public distribution pipes because of very inhomogeneous ownership and responsibility, which seems to be only limited by the creativity of the utility statutes that define them (at least in Switzerland). More and more utilities start registering the properties and failures of household connections, which provide the basis for better addressing this upcoming challenge in the future.

- The MCDA about pipe rehabilitation strategies did only consider combinations of repair and replacement. With increasing practical experience about the durability, quality implications, and costs of different pipe renovation strategies available, it would be interesting to additionally include renovation into the comparison of rehabilitation strategies.

- Part of the dimensioning and predictions underlying the modeling in this study are based on across-the-board assumptions and gross simplifications of the distribution system. Given the high importance of supply system reliability, and water quality, more efforts should be spent on better predictions of the respective attributes. With more time (and GIS data) available, it would be very beneficial to set up and use a hydraulic model of the water system to get a better understanding of the performance of different alternatives on water pressure, the extent of failures and potential third-party damage, water losses, and service interruptions. It goes without saying, that the predictions of other attributes can, and probably should, also be improved.

- To derive recommendations as to which uncertain parameters have the highest impact on the results of the MAUT analysis and whether additional efforts should be spent on improving attribute predictions rather than more detailed preference elicitations, an extended global sensitivity analysis would be necessary. An extended sensitivity analysis which includes both the uncertain attribute parameters and preference parameters will be helpful to indicate the predictions of which attribute should be improved, and more research and / or modeling efforts be spent on. Such an analysis would not only support the understanding of the modeling and interpretation of the results of this study. Global sensitivity analysis are still rare in the field and to my knowledge limited to very simple examples (e.g. analysis of the uncertainty of a small number of weights and attributes, see Gómez Delgado and Bosque Sendra, 2004; Mustajoki et al., 2006; Saltelli et al., 1999a).

Other issues and research potential

- Existing performance indicators (Alegre et al., 2006) form a well-accepted basis for water infrastructure decision support. This can be built on by re-formulating them as fundamental objectives and attributes for use in formal decision analysis.

- There is an increasing demand to integrate the management of municipal infrastructure networks, but few models actually consider this demand. As many of the water rehabilitation projects are actually triggered by other ground works (e.g. wastewater, gas, roads), synergy effects of combined modeling and asset management should receive more focus in the future.

Practicability of approaches

As already mentioned, I did not address the challenge of limited institutional, financial, and professional capacities. Any recommendation on how to overcome these would be highly speculative and requires a sound understanding of water governance, policy, and financing, among other fields. My following statements are therefore limited to some thoughts about the developed tools, and potential simplifications to make them more practicable. I thereby borrow from discussions with practitioners during my work in the NRP 61 research project "Sustainable Water Infrastructure Planning", exchanges with other scholars or members of the Swiss Gas and Water Industry Association (SVGW) that are not documented elsewhere.

(1) Software and models

a. Contrary to the wishes of many practitioners, this work did not result in the development of any off-the-shelf software. While I trust that any of the developed or adapted models underlying this thesis can be efficiently implemented into a useable and attractive software, the necessary expertise and time efforts for their understanding will probably exceed what can be requested from small water utilities. There might be more potential in training experts (e.g. consultants) to apply the developed approaches whether or not implemented into specific software.

b. Making the Bayesian parameter inference procedure accessible to a wider public could be beneficial to support the assessment of condition and future failures in small networks. This should be possible based on the available R codes, but requires training.

c. The R package "utility" (Reichert et al., 2013) used for implementing the preference model and MAUT evaluations is very flexible (not limited to "usual" simplification assumptions, chapter 6) and is suitable to propagate uncertainties. Even so, setting up the model and performing calculations requires time and some programming skills, which might be a hindrance for MCDA practitioners (but can probably be overcome with specific training).

d. A small R program (also based on the utility package) used to calculate the (additive) weights of objectives and the trade-offs for consistency checks was very helpful during elicitation. Its further development towards an interactive support tool for preference elicitation would be beneficial.

(2) Simplifying the approach

a. The presented MCDA study for 'good water supply infrastructure' is much more complex than what practitioners look out to. The efforts for problem structuring appear overwhelming to outsiders, not to speak of the objectives hierarchy (44 objectives), the many characteristics of the alternatives (17), the prediction of attributes under four scenarios (i.e. 120) , and on top of that the uncertainty analysis. This is too far away from anything that could be judged a well communicable example. It furthermore addressed a created decision problem that the participants were not compelled to solve. I trust, if we work on smaller, real, and more easily communicable decision problems, we will be more likely to convince practice that this approach is helpful to address upcoming infrastructure challenges.

b. The “small” MCDA (on finding a ‘good rehabilitation strategy’) might be more suitable and could be adapted (e.g. interviewing asset managers, using only a selection of the alternatives) so as to serve as a better example for demonstration.

c. Similarly, the approach for failure and rehabilitation modeling should be much simplified to create more communicable examples. The obtained priors can be used as a basis for the estimation of condition, failure rates and rehabilitation demand in small utilities without failure records and also for the communication of uncertainty. Using only a selection of rehabilitation alternatives, the necessary financial efforts and benefits can be presented.

d. The findings obtained with this approach can be furthermore used to demonstrate that thorough data keeping is of direct practical value even in small utilities and that sound failure modeling and strategic asset management is possible in small water utilities.

Acknowledgements

First and foremost I want to thank my supervisors Judit Lienert, Max Maurer, and Peter Reichert for their outstanding guidance and support throughout the development of this thesis. Your constant belief in me and backing of my work has been a major source of motivation. I also thank Eberhard Morgenroth for his support, particularly during the initial phases of this dissertation.

This work has greatly benefitted from collaborations with many people. On the science side, I would like to especially thank Nele Schuwirth and Andreas Scheidegger for their great commitment and tutoring in modeling (and debugging) in R, applied statistics, or MCDA elicitation techniques. I further thank Christoph Egger, Cristina Fritzsche, Florian Schnetzer, Jun Zheng, and Anja Zorn for their good collaboration, engagement, and support within the SWIP project. Adelaida Arreaza, Markus Günther, and Katrin Wolter prepared their master and bachelor theses under my supervision – thank you for your steady engagement, interesting discussions, and good work! Also not to be forgotten are the volunteers for testing the online survey and MCDA guideline. Unless already mentioned, I thank Thomas Hug, Mira Kattwinkel, Eva Lieberherr, and Hans-Peter Zöllig for assisting me in this task.

On the practice side, my special thanks go to Sébastien Apothéloz, Markus Biner, Christoph Meyer, Andreas Peter, Adrian Rieder, and Thomas Weyermann. Your interest, practical experience, insights, and the provided data have fundamentally shaped this work.

In addition, I extend my gratitude to all those practitioners and scientists who supported me and the SWIP project with expertise, participation in workshops, online surveys, and / or interviews, namely (unless already mentioned): Thomas Bodmer, Dieter Börlin, Hugo Brändle, Matthias Freiburghaus, Daniela Fuchs-Hanusch, Heiko Gebauer, Ruedi Grob, Frederik Hammes, Daniel Hartmann, Jürg Haas, Thomas Hillenbrand, Juliane Hollender, Mark Honti, Markus Hugger, Daniel Hunkeler, Fritz Jäckli, Urs Kamm, Roland Käser, Stefan Koetzsch, Ivana Logar, Christoph Meier, Ernst Meili, Christian Moeck, Kurt Nyffenegger, Andreas Peter, Rainer Rhunke, Jörg Rieckermann, Paul Ruckstuhl, Robert Rupp, Mario Schirmer, Jürg Schmid, Stefan Schmid, Hans-Ulrich Stauber, Kurt Steiner, Stefan Vollenweider, Jan Vreeburg, Urs von Gunten.

I am very grateful to my former and current colleagues at the Siam / SWW / ESS departments, to Raoul Schaffner, "our front ladies" at the Eawag reception, the Lib4ari staff, IT supporters, and the Ipazia cluster administrators for the very friendly and collaborative atmosphere.

Lastly, I would like to express my heartfelt thanks to my friends, partner, and family for their everlasting support under all circumstances.

This research was conducted within the research project 'Sustainable Water Infrastructure Planning (SWIP)'(project number 406140_125901/1) as part of the National Research Program 61 on Sustainable Water Management and funded by the Swiss National Science Foundation (SNF).

References

Abdellaoui, M., Bleichrodt, H., L'Haridon, O., 2008. A tractable method to measure utility and loss aversion under prospect theory. Journal of Risk and Uncertainty 36(3) 245-266.

Abdellaoui, M., Bleichrodt, H., Paraschiv, C., 2007. Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. Management Science 53(10) 1659-1674.

Afify, A., 2010. Prioritizing desalination strategies using multi-criteria decision analysis. Desalination 250(3) 928-935.

AfU, 2006. Generelle Wasserversorgungsplanung (GWP)- Richtlinie zur Erstellung der GWP (General water supply plan (GWP)- guideline for the preparation of the GWP), In: Solothurn), A.f.U.O.f.t.E.o.t.c.o. (Ed.).

AGWAM, 2013. Increasing water scarcity, also for Swiss agriculture. Project of the National Research Program NRP 61. http://www.nfp61.ch/E/projects/cluster-water-management/water_agriculture/Pages/default.aspx. Accessed 06 December 2013

Alegre, H., 2010. Is strategic asset management applicable to small and medium utilities? Water Science and Technology 62(9) 2051-2058.

Alegre, H., Baptista, J.M., Cabrera, E., Cubillo, F., Duarte, W., Hirner, W., Merkel, W., Parena, R., 2006. Performance Indicators for Water Supply Services, 2nd ed. IWA International Water Association.

Allais, M., 1953. Le comportement de l'homme rationel devant le risque: Critique des postulats et axiomes de l'ecole américaine. Econometrica 21(4) 503-546.

Alpert, M., Raiffa, H., 1982. A progress report on the training of probability assessors, In: Kahneman, D., Slovic, S., Tversky, A. (Eds.), Judgment and Uncertainty: Heuristics and Biases. Cambridge University Press, pp. 294-305.

Alvisi, S., Franchini, M., 2010. Comparative analysis of two probabilistic pipe breakage models applied to a real water distribution system. Civil Engineering and Environmental Systems 27(1) 1-22.

Ana, E., Bauwens, W., Broers, O., 2009. Quantifying uncertainty using robustness analysis in the application of ORESTE to sewer rehabilitation projects prioritization—Brussels case study. Journal of Multi-Criteria Decision Analysis 16(3-4) 111-124.

Ananda, J., Herath, G., 2005. Evaluating public risk preferences in forest land-use choices using multi-attribute utility theory. Ecological Economics 55(3) 408-419.

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(10) 2535-2548.

Arnbjerg-Nielsen K, H.S., F., 2009. Feasible adaptation strategies for increased risk of flooding in cities due to climate change. Water Science and Technology 60(2) 273-281.

Arreaza, A.S. (2011) Robust Elicitation of Imprecise Expert Knowledge for Water Infrastructure Management. MSc thesis, ETH Zurich, Zürich, Switzerland.

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(3) 200-209.

AUE, 2012. Wegleitung für Generelle Wasserversorgungsprojekte (Guidance for water supply projects), In: Basel-Land), A.f.U.u.E.O.f.e.p.a.e.o.t.c. (Ed.).

AWA, 2011. Generelle Wasserversorgungsplanung GWP- Wegleitung 2011, In: Berne), A.A.f.W.u.A.d.K.B.O.f.W.a.W.o.t.c.o. (Ed.).

AWEL, 2003. Phosphorbelastung des Greifensees. (Phosphorous pullution of Greifensee lake). Amt für Abfall, Wasser, Energie und Luft (Office for waste, water, energie, and air, ct. Zürich). http://www.awel.zh.ch/internet/baudirektion/awel/de/wasserwirtschaft/gewaesserqualitaet/seen_kenngroessen.html. Accessed 21 June 2012

AWEL, 2006. Zustand der Fliessgewässer in den Einzugsgebieten von Glatt und Greifensee - Messkampagne 2004/ 2005. Amt für Abfall, Wasser, Energie und Luft, Kt. Zürich. http://www.awel.zh.ch/internet/baudirektion/awel/de/wasserwirtschaft/veroeffentlichungen.html. Accessed 15 November 2012

AWEL, 2010. Richtlinien für die Erstellung von Generellen Wasserversorgungsprojekten (GWP) (Guideline for the preparation of a general water supply project), In: AWEL Amt für Abfall, W., Energie und Luft (Office for Waste, Water, Energie and Air of the canton of Zurich) (Ed.).

AWEL, 2013. Grundwasserqualität - Messdaten. http://www.awel.zh.ch/internet/baudirektion/awel/de/wasserwirtschaft/grundwasser/messdaten_gw/gw_quali.html#a-content. Accessed 18 August 2013

AWWA, 2001. Reinvesting in Drinking Water Infrastructure. Dawn of the Replacement Era. American Water Works Association. http://www.win-water.org/reports/infrastructure.pdf. Accessed 23 August 2013

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

Balkema, A.J., Preisig, H.A., Otterpohl, R., Lambert, A.J.D., Weijers, S.R., 2001. Developing a model based decision support tool for the identification of sustainable treatment options for domestic wastewater. Water Science and Technology 43(7) 265-270.

Baur, R., Le Gauffre, P., Saegrov, S., 2003. Multi-criteria decision support for annual rehabilitation programmes in drinking water networks. Water Science and Technology: Water Supply 3(1) 43-50.

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: Ehrgott, M. (Ed.), Trends in Multiple Criteria Decision Analysis, pp. 209 - 239.

Berger, J.O., 1990. Robust Bayesian Analysis- Sensitivity to the prior. Journal of Statistical Planning and Inference 25(3) 303-328.

Bernardo, J.M., Smith, A.F.M., 2000. Bayesian Theory, 1st Edition ed. Wiley.

BFS, 2010. Insitutionelle Gliederungen. http://www.bfs.admin.ch/bfs/portal/de/index/regionen/11/geo/institutionelle_gliederungen/01b.html. 17 August 2010

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(11) 1498-1514.

Bond, S.D., Carlson, K.A., Keeney, R.L., 2008. Generating objectives: Can decision makers articulate what they want? Management Science 54(1) 56-70.

Bond, S.D., Carlson, K.A., Keeney, R.L., 2010. Improving the Generation of Decision Objectives. Decision Analysis 7(3) 238-255.

Borcherding, K., Eppel, T., Detlof von, W., 1991. Comparison of Weighting Judgements in Multiattribute Utility Measurement. Management Science 37(12) 1603-1619.

Borsuk, M.E., Maurer, M., Lienert, J., Larsen, T.A., 2008. Charting a path for innovative toilet technology using multicriteria decision analysis. Environmental Science and Technology 42(6) 1855-1862.

Bosshard, T., Kotlarski, S., Ewen, T., Schär, C., 2011. Spectral representation of the annual cycle in the climate change signal. Hydrology and Earth System Sciences 15(9) 2777-2788.

Bousquet, N., 2008. Diagnostics of prior-data agreement in applied Bayesian analysis. Journal of Applied Statistics 35(9) 1011-1029.

Boxall, J.B., O'Hagan, A., Pooladsaz, S., Saul, A.J., Unwin, D.M., 2007. Estimation of burst rates in water distribution mains. Proceedings of the Institution of Civil Engineers- Water Management 160 73-82.

Bradfield, R., Wright, G., Burt, G., Cairns, G., Van Der Heijden, K., 2005. The origins and evolution of scenario techniques in long range business planning. Futures 37(8) 795-812.

Brown, R.R., Farrelly, M.A., 2009. Delivering sustainable urban water management: A review of the hurdles we face. Water Science and Technology 59(5) 839-846.

Brown, R.R., Keath, N., Wong, T.H.F., 2009. Urban water management in cities: historical, current and future regimes. Water Science and Technology 59(5) 847-855.

Bundi, U., Peter, A., Frutiger, A., Hütte, M., Liechti, P., Sieber, U., 2000. Scientific base and modular concept for comprehensive assessment of streams in Switzerland. Hydrobiologia (422/423) 477-487.

Burn, S., Tucker, S., Rahilly, M., Davis, P., Jarrett, R., Po, M., 2003. Asset planning for water reticulation systems - The PARMS model. Water Science and Technology: Water Supply 3(1-2) 55-62.

Burns, P., Hope, D., Roorda, J., 1999. Managing infrastructure for the next generation. Automation in construction 8(6) 689-703.

Butler, J., Jia, J., Dyer, J., 1997. Simulation techniques for the sensitivity analysis of multi-criteria decision models. European Journal of Operational Research 103(3) 531-546.

Butler, J., McEntee, B., Onof, C., Hagger, A., 2007. Sewer storage tank performance under climate change. Water Science and Technology 56(12) 29-35.

Calizaya, A., Meixner, O., Bengtsson, L., Berndtsson, R., 2010. Multi-criteria decision analysis (MCDA) for integrated water resources management (IWRM) in the Lake Poopo basin, Bolivia. Water Resources Management 24(10) 2267-2289.

Cardoso, M.A., Santos Silva, M., Coelho, S.T., Almeida, M.C., Covas, D.I.C., 2012. Urban water infrastructure asset management - A structured approach in four water utilities. Water Science and Technology 66(12) 2702-2711.

Carriço, N., Covas, D.I.C., Céu Almeida, J.P., Leitão, J.P., Alegre, H., 2012. Prioritization of rehabilitation interventions for urban water assets using multiple criteria decision-aid methods. Water Science and Technology 66(5) 1007-1014.

Carrión, A., Solano, H., Gamiz, M.L., Debón, A., 2010. Evaluation of the Reliability of a Water Supply Network from Right-Censored and Left-Truncated Break Data. Water Resources Management 24(12) 2917-2935.

Cashman, A., Ashley, R., 2008. Costing the long-term demand for water sector infrastructure. Foresight 10(3) 9-26.

CH2011, 2011. Swiss Climate Change Scenarios 2011. Zurich, Switzerland, ISBN: 978-3-033-03065-7.

Chateauneuf, A., Cohen, M., 2000. Choquet Expected Utility Model: A New Approach to Individual Behavior under Uncertainty and Social Welfare, In: Grabisch, M., Murofushi, T., Sugeno, M. (Eds.), Fuzzy Measures and Integrals: Theory and Applications. Physica-Verlag: Berlin, pp. 289-313.

Chen, Z., Ngo, H.H., Guo, W.S., Listowski, A., O'Halloran, K., Thompson, M., Muthukaruppan, M., 2012. Multi-criteria analysis towards the new end use of recycled water for household laundry: A case study in Sydney. Science of the Total Environment (438) 59-65.

Christodoulou, S., Charalambous, C., Adamou, A., 2008. Rehabilitation and maintenance of water distribution network assets. Water Science and Technology: Water Supply 8(2) 231-237.

Clemen, R.T., Reilly, T., 2001. Making Hard Decisions, 2nd ed. Duxbury.

Clemen, R.T., Winkler, R.L., 1999. Combining Probability Distributions From Experts in Risk Analysis. Risk Analysis 19(2) 187-203.

Cobb, C.W., Douglas, P.H., 1928. A theory of production. The American Economic Review 18(1) 139-165.

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

Cooke, R.M., Goossens, L.L.H.J., 2008. TU Delft expert judgment data base. Reliability Engineering & System Safety 93(5) 657-674.

Coolen, F.P.A., 1996. On Bayesian reliability analysis with informative priors and censoring. Reliability Engineering and System Safety 53(1) 91-98.

Cordazzo, V., 2006. Die Sterblickkeit der Schweizer Geburtenjahrgänge 1900 bis 2030, Demos. Informationen aus der Demografie. Federal Statistical Office (BFS): Neuchâtel, p. 32.

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., Truffer, B., Gujer, W., 2011. Tackling uncertainty in infrastructure sectors through strategic planning: the contribution of discursive approaches in the urban water sector. Water Policy 13(3) 299-316.

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(2) 30-50.

Dong, C., Schoups, G., van de Giesen, N., 2013. Scenario development for water resource planning and management: A review. Technological Forecasting and Social Change 80(4) 749-761.

Dridi, L., Mailhot, A., Parizeau, M., Villeneuve, J.-P., 2009. Multiobjective Approach for Pipe Replacement Based on Bayesian Inference of Break Model Parameters. Journal of Water Resources Planning and Management 135(5) 344-354.

Duckstein, L., Treichel, W., Elmagnouni, S., 1994. Ranking groundwater management alternatives by multicriteriopn analysis. Journal of Water Resources Planning and Management-Asce 120(4) 546-565.

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(2) 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(4) 456-464.

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

Dyer, J.S., Sarin, R.K., 1979. Measurable Multiattribute Value Functions. Operations Research 27(4) 810-822.

Dyer, J.S., Sarin, R.K., 1982. Relative Risk Aversion. Management Science 28(8) 875-886.

Eawag, 2009. Wasserversorgung 2025- Vorprojekt. Wissensstand zur Wasserversorgung in der Schweiz. Eawag. FOEN, Federal Office for the Environment. http://www.bafu.admin.ch/wassernutzung/02792/index.html?lang=en. Accessed 16 October 2010

Economides, J.A., 2012. Public outreach: An engineer's perspective. Journal - American Water Works Association 104(1) 74-76.

Economou, T., Kapelan, Z., Bailey, T., 2009. A Zero-Inflated Bayesian Model for the Prediction of Water Pipe Bursts. Geotechnical Special Publication(187) 724-734.

Eder, G., Duckstein, L., Nachtnebel, H.P., 1997. Ranking water resource projects and evaluating criteria by multicriterion Q-analysis: an Austrian case study. Journal of Multi-Criteria Decision Analysis 6(5) 259-271.

EFQM, 2012. European Foundation for Quality Management. http://www.efqm.org/en. Accessed 6 November 2012

Egger, C., Scheidegger, A., Reichert, P., Maurer, M., 2013. Sewer deterioration modelling with data lacking historical records Water Research(47) 6762-6779.

Eisenbeis, P., Rostum, J., Le Gat, Y., 1999. Statistical models for assessing the technical state of water networks: some european experiences, AWWA American Water Works Association Annual Conference. AWWA: Chicago, Illinois.

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(4) 643-669.

Engelhardt, M., Savic, D., Skipworth, P., Cashman, A., Saul, A., Walters, G., 2003. Whole life costing: Application to water distribution network. Water Science and Technology: Water Supply 3(1-2) 87-93.

Engelhardt, M.O., Skipworth, P.J., Savic, D.A., Saul, A.J., Walters, G.A., 2000. Rehabilitation strategies for water distribution networks: a literature review with a UK perspective. Urban Water 2(2) 153-170.

Etter, B., Kohn, T., 2007. Struvite recovery from urine at community scale in Nepal. Intermediate project report submitted to EPFL. Eawag. Dübendorf. Switzerland.

Farmani, R., Walters, G.A., Savic, D.A., 2005. Trade-off Between Total Cost and Reliability for Anytown Water Distribution Network. Journal of Water Resources Planning and Management 131(3) 161-171.

Fatichi, S., Ivanov, V., Caporali, E., 2011. AWE-GEN: Advanced weather generator technical reference version 1.0. University of Michigan, Ann Arbor, Michigan, USA.

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.

Fichtner Asset Services & Technologies (2013) FAST-Fichtner Asset Services & Technologies. v13.1. Solution partnership with Vesta Business Simulations. http://www.optnet.de/index.php?id=135, Accessed 08 April 2013.

Figueira, J., Greco, S., Ehrgott, M., 2005. Multiple Criteria Decision Analysis: State of the Art Surveys. Springer Science + Business Media, Inc., Boston.

FIV, 2009. Verordnung des EDI über Fremd- und Inhaltsstoffe in Lebensmitteln (Fremd- und Inhaltsstoffverordnung, FIV) vom 26. Juni 1995 (Stand am 1. Oktober 2009), In: Inneren, E.D.d. (Ed.).

FOEN, 2010. Methoden zur Untersuchung und Beurteilung der Fliessgewässer; chemisch-physikalische Erhebungen, Nährstoffe. (Methods for the investigation and assessment of watercourses; chemico-physical parameters, nutrients). Swiss Federal Office for the Environment. http://www.umwelt-schweiz.ch/uv-1005-d. Accessed 15 November 2012

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

Frederick, S., Loewenstein, G., O'Donoghue, T., 2002. Time discounting and time preference: A critical review. Journal of Economic Literature 40(2) 351-401.

French, S., 1995. Uncertainty and Imprecision: Modelling and Analysis. Journal of the Operational Research Society 46(1) 70-79.

French, S., 2003. Modelling, Making Inferences and Making Decision: The Roles of Sensitivity Analysis. Sociedad de Estadística e Investigación Operativa (TOP) 11(2) 229-251.

Frühwirth-Schnatter, S., 2006. Finite mixture and markov switching models. Springer, New York.

Fuchs-Hanusch, D., Gangl, G., Kornberger, B., Kölbl, J., Hofrichter, J., Kainz, H., 2008. PiReM- Pipe Rehabilitation Management. Developing A Decision Support System for Rehabilitation Planning of Water Mains. Water Practice & Technology 3(1) 1-9.

Fuchs-Hanusch, D., Kornberger, B., Friedl, F., Scheucher, R., 2012. Whole of Life Cost Calculations for Water Supply Pipes. Water Asset Management International 8(2) 19-24.

Fuchs, D. (2001) Decision Support Systeme für die Rehabilitationsplanung von Wasserrohrnetzen. PhD thesis, Graz University of Technology, Graz, Austria.

Gander, B., 2009. Klimaänderung und Wasserversorgungen- Information und Anpassungsstrategien. GWA(4/2009) 241-249.

Gangl, G. (2008) Rehabilitationsplanung von Trinkwassernetzen. PhD thesis, Graz University of Technology, Graz, Austria.

Ganoulis, J., 2003. Evaluating alternative strategies for wastewater recycling and reuse in the Mediterranean area. Water Science and Technology: Water Supply 3(4) 11-19.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian data analysis, 2nd ed. Chapmann & Hall/CRC.

Gelman, A., Hill, J., 2009. Data analysis using regression and multilevel/ hierarchical models. Cambridge University Press, New York.

Genest, C., Zidek, J.V., 1986. Combining Probability Distributions: A Critique and an Annotated Bibliography. Statistical Science 1(1) 114-135.

Gillies, D., 1991. Intersubjective probability and confirmation theory. British Journal for the Philosophy of Science 42(4) 513-533.

Giustolisi, O., Laucelli, D., Savic, D.A., 2006. Development of rehabilitation plans for water mains replacement considering risk and cost-benefit assessment. Civil Engineering and Environmental Systems 23(3) 175-190.

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(6) 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(4) 449-466.

Goodwin, P., Wright, G., 2001. Enhancing strategy evaluation in scenario planning: A role for decision analysis. Journal of Management Studies 38(1) 1-16.

Graves, S.B., Ringuest, J.L., 2009. Probabilistic dominance criteria for comparing uncertain alternatives: A tutorial. Omega(37) 346-357.

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., Ohlson, D., 2012a. Structured Decision Making- A practical guide to Environmental Management Choices, 1st ed. Wiley-Blackwell.

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T., Ohlson, D., 2012b. When experts disagree (and better science won't help much): Using structured deliberations to support endangered species recovery planning. Journal of Environmental Management (105) 30-43.

GSchV, 2011. Gewässerschutzverordnung vom 28. Oktober 1998 (Stand am 1. August 2011).

Guest, J.S., Skerlos, S.J., Barnard, J.L., Beck, M.B., Daigger, G.T., Hilger, H., Jackson, S.J., Karvazy, K., Kelly, L., Macpherson, L., Mihelcic, J.R., Pramanik, A., Raskin, L., van Loosdrecht, M.C.M., Yeh, D., Love, N.G., 2009. A new planning and design paradigm to achieve sustainable resource recovery from wastewater. Environmental Science & Technology 43(16) 6126-2130.

Gujer, W., 2002. Siedlungswasserwirtschaft, 3 ed. Springer Verlag, Berlin.

Gustafson, J.-M., Clancy, D.V., 1999. Modeling the occurrence of breaks in cast iron water mains using methods of survival analysis, AWWA American Water Works Assiciation Annual Conference. AWWA: Chicago, Illinois.

GVZ, 2011. Richtlinien für die Ausführung der Löschwasserversorgung und die Subventionen der Gebäudeversicherung Kanton Zürich (GVZ) an Hydranten, Hydrantenkontrollwartung und Hydrantenunterhalt, In: Zürich, G.K. (Ed.). GVZ Zürich: Zürich.

Hadar, J., Russell, W.R., 1969. Rules for Ordering Uncertain Prospects. The American Economic Review 59(1) 25-34.

Haffejee, M., Brent, A.C., 2008. Evaluation of an integrated asset life-cycle management (ALCM) model and assessment of practices in the water utility sector. Water SA 34(2) 285-290.

Hajkowicz, S., 2008. A comparison of multiple criteria analysis techniques for water resource management. European Journal of Operational Research (184) 255-265.

Hajkowicz, S., Collins, K., 2007. A review of multiple criteria analysis for water resource planning and management. Water Resources Management 21(9) 1553-1566.

Hämälainen, R.P., Kettunen, E., Marttunen, M., Ehtamo, H., 2001. Evaluating a framework for multi-stakeholder decision support in water resources management. Group Decision and Negotiation 10(4) 331-353.

Hanoch, G., Levy, H., 1969. The Efficiency Analysis of Choices Involving Risk. The Review of Economic Studies 36(3) 335-346.

Harrell, J., Frank E., 2001. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. Springer- Verlag.

Heather, A.I.J., Bridgeman, J., 2007. Water industry asset management: a proposed service-performance model for investment. Water and Environment Journal(21) 127-132.

Herlyn, A., Maurer, M., 2007. Zustand und Investitionsbedarf der Schweizer Abwasserentsorgung. Schweizer Gemeinde 11/07 14-17.

Hershey, J.C., Schoemaker, P.J.H., 1985. Probability versus certainty equivalence methods in utility measurement- are they equivalent? Management Science 31(10) 1213-1231.

Herz, R.K., 1995. Alterung und Erneuerung von Infrastrukturbeständen- Ein Kohortenüberlebensmodell (in German: Ageing and renewal of infrastructure- a cohort survival model), Jahrbuch für Regionalwissenschaft. Institut für Städtebau und Landesplanung, Universität Karlsruhe, pp. 7-29.

Herz, R.K., 1998. Exploring rehabilitation needs and strategies for water distribution networks. Journal of Water Supply Research and Technology AQUA 47(6) 275.

Howard, R.A., 1988. Decision analysis- Practice and promise. Management Science 34(6) 679-695.

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(19) 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(6) 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(4) 278-290.

Hyde, K.M., May, R., Leak, M., 2011. Analysis of options to increase supply for the Virginia pipeline irrigation scheme utilising Multi-criteria decision analysis. Water Practice and Technology 6(1).

iWaQa, 2013. Integrated river water quality management. Project of the National Research Program NRP 61. http://www.nfp61.ch/E/projects/cluster-water-management/integrated-river_water-quality_management/Pages/default.aspx. Accessed 06 December 2013

Jacobi, S.K., Hobbs, B.F., 2007. Quantifying and Mitigating the Splitting Bias and Other Value Tree-induced Weighting Biases. Decision Analysis 4(4) 194-210.

Jafar, R., Shahrour, I., Juran, I., 2010. Application of artificial neural networks (ANN) to model the failure of urban water mains. Mathematical and Computer Modelling 51(9-10) 1170-1180.

Janis, I., 1972. Vistims of Groupthink. Houghton Mifflin, Boston, MA.

Janis, I., 1982. Groupthink: Psychological Studies of Policy Decisions and Fiascoes, 2nd ed. Houghton Mifflin, Boston, MA.

Jessop, A., 2011. Using imprecise estimates for weights. Journal of the Operational Research Society 62(6) 1048-1055.

Jiménez, A., Ríos-Insua, S., Mateos, A., 2006. A generic multi-attribute analysis system. Computers & Operations Research 33(4) 1081-1101.

Jones, D.S., Efroymson, R.A., Armstrong, A.Q., Muhlheim, M.D., Carnes, S.A., 2004. Integrated Risk Assessment for Individual Onsite Wastewater Treatment Systems. St. Louis, Oak Ridge National Laboratory. Prepared for the National Decentralized Water Resources Capacity Development Project, Project No. WU-HT-01-18.

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(12) 17-25.

Jouini, M.N., Clemen, R.T., 1996. Copula models for aggregating expert opinions. Operations Research 44(3) 444-457.

Kahneman, D., Tversky, A., 1979. Prospect Theory: An Analysis of Decision under Risk. Econometrica 47(2) 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 (6) 169-188.

Kappeler, J., Neubert, K., Woll, L., 2010. Benchmarking für Wasserversorgungen: Erkenntnisse für die Branche. GWA 90(4) 305-314.

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(5-6) 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(5-6) 125-137.

Karvetski, C.W., Lambert, J.H., Linkov, I., Keisler, J., 2011. Integration of decision analysis and scenario planning for coastal engineering and climate change. IEEE Transactions on Systems, Man, and Cybernetics 41(1) 63-73.

Katsoyiannis, I.A., Canonica, S., von Gunten, U., 2011. Efficiency and energy requirements for the transformation of organic micropollutants by ozone, O3/H2O2 and UV/H2O2. Water Research 45(13) 3811-3822.

Keeney, R.L., 1982. Decision Analysis: An Overview. Operations Research 30(5) 803-838.

Keeney, R.L., 1992. Value-focused thinking- a path to creative decisionmaking, 1 ed. Harvard University Press, Cambridge.

Keeney, R.L., McDaniels, T., Ridge-Cooney, V.L., 1996. Using values in planning wastewater facilities for metropolitan Seattle. Water Resources Bulletin 32(2) 293-303.

Keeney, R.L., Raiffa, H., 1976. Decisions with Multiple Objectives: Preferences and Value Tradeoffs. John Wiley and Sons, New York.

Keeney, R.L., Raiffa, H., 1993. Decisions with multiple objectives. Preferences and Value tradeoffs, 2 ed. Cambridge University Press, Cambridge.

Kendall, M.G., 1938. A new measure of rank correlation. Biometrika 30(1-2) 81-93.

Kerr, N.L., Tindale, R.S., 2004. Group performance and decision making. Annual Review of Psychology (55) 623-655.

Kilsby, C.G., Jones, P.D., Burton, A., Ford, A.C., Fowler, H.J., Harpham, C., James, P., Smith, A., Wilby, R.L., 2007. A daily weather generator for use in climate change studies. Environmental Modelling & Software 22(12) 1705-1719.

Kiureghian, A.D., Didevsen, O., 2009. Aleatory or epistemic? Does it matter? Structural Safety 31(2) 105-112.

Klein, J.P., Moeschberger, M.L., 2003. Survival analysis: techniques for censored and truncated data, 2nd ed. New York: Springer, New York.

Kleiner, Y., Nafi, A., Rajani, B., 2009. Planning Renewal of Water Mains While Considering Deterioration, Economies of Scale and Adjacent Infrastructure, 2nd International Conference on Water Economics, Statistics and Finance. IWA: Alexandroupolis, Greece.

Kleiner, Y., Rajani, B., 1999. Using Limited Data to Assess Future Needs. Journal / AWWA 91(7) 47-61.

Kleiner, Y., Rajani, B., 2001. Comprehensive Review of structural deterioration of water mains: statistical models. Urban Water(3) 131-150.

Kleiner, Y., Rajani, B., 2004. Quantifying Effectiveness of Cathodic Protection in Water Mains: Theory. Journal of Infrastructure Systems 10(2) 43-51.

Kleiner, Y., Rajani, B., 2010. I-WARP: Individual Water Main Renewal Planner. Drinking Water Engineering and Science 3(1) 71-77.

Kodikara, P.N., Perera, B.J.C., Kularathna, M.D.U.P., 2010. Stakeholder preference elicitation and modelling in multi-criteria decision analysis- A case study on urban water supply. European Journal of Operational Research(206) 209-210.

Köksalan, M., Wallenius, J., Zionts, S., 2013. An Early History of Multiple Criteria Decision Making. Journal of Multi-Criteria Decision Analysis 20(1-2) 87-94.

Korving, H., van Noortwijk, J.M., 2008. Bayesian updating of a prediction model for sewer degradation. Urban Water Journal 5(1) 51-57.

Kropp, I., Baur, R., 2005. Integrated failure forecasting model for the strategic rehabilitation planning process. Water Science and Technology: Water Supply 5(2) 1-8.

Krueger, T., Page, T., Hubacek, K., Smith, L., Hiscock, K., 2012. The role of expert opinion in environmental modelling. Environmental Modelling & Software 36(0) 4-18.

Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecology Letters 13(7) 900-914.

Kynn, M., 2008. The 'heuristics and biases' bias in expert elicitation. Journal of the Royal Statistical Society: Series A (Statistics in Society) 171(1) 239-264.

Kysely, J., Gaal, L., Beranova, R., Plavcova, E., 2011. Climate change scenarios of precipitation extremes in Central Europe from ENSEMBLES regional climate models. Theoretical and Applied Climatology 104(3-4) 529-542.

Lahdelma, R., Hokkanen, J., Salminen, P., 1998. SMAA - Stochastic multiobjective acceptability analysis. European Journal of Operational Research 106(1) 137-143.

Lahdelma, R., Salminen, P., 2012. The shape of the utility or value function in stochastic multicriteria acceptability analysis. OR Spectrum 34(4) 785-802.

Lambert, J., Karvetski, C.W., Spencer, D., Sotirin, B., Liberi, D., Zaghloul, H., Koogler, J., Hunter, S., Goran, W., Ditmet, R., Linkov, I., 2012. Prioritizing infrastructure investments in Afghanistan with multiagency stakeholders and deep uncertainty of emergent conditions. Journal of Infrastructure Systems 18(2) 155-166.

Lang, D.J., Wiek, A., Bergmann, M., Stauffacger, M., Martens, P., Moll, P., Swilling, M., Thomas, C.J., 2012. Transdisciplinary research in sustainability science: practice, principles, and challenges. Sustainability Science 7(1) 25-43.

Langhans, S.D., Lienert, J., Schuwirth, N., Reichert, P., 2013. How to make river assessments comparable: A demonstration for hydromorphology. Ecological Indicators (32) 264-275.

Langhans, S.D., Reichert, P., 2011. Einbettung von Verfahren zur Fliessgewässerbewertung in ein übergeordnetes Gewässermanagementkonzept; Vorschläge am Beispiel des Modulstufenkonzepts (in German: Embedding procedures for stream assessment into an overriding water management concept; propositions using the example of the Swiss Modular Concept). Wasser Energie Luft 103(3) 204-214.

Langhans, S.D., Reichert, P., Schuwirth, N., submitted. The method matters: indicator aggregation in ecological assessment.

Larrick, R.P., Soll, J.B., 2006. Intuitions about combining opinions: Misappreciation of the averaging principle. Management Science 52(1) 111-127.

Larsen, T.A., Alder, A.C., Eggen, R.I.L., Mauer, M., Lienert, J., 2009. Source separation: Will we see a paradigm shift in wastewater handling? Environmental Science & Technology 43(16) 6121-6125.

Larsen, T.A., Udert, K.M., Lienert, J., 2012. Source Separation and Decentralization for Wastewater Management. IWA Publishing: London, New York.

Lautenschlager, K., Boon, N., Wang, Y., Egli, T., Hammes, F., 2010. Overnight stagnation of drinking water in household taps induces microbial growth and changes in community composition. Water Research 44(17) 4868-4877.

Le Gat, Y. (2009) Une extension du processus de Yule pour la modélisation stochastique des événements récurrents- Application aux défaillances de canalisations d'eau sous pression. PhD thesis., ENGREF Paris.

Lele, S.R., Allen, K.L., 2006. On using expert opinion in ecological analyses: a frequentist approach. Environmetrics 17(7) 683-704.

Libralato, G., Volpi, G.A., Avezzù, F., 2012a. To centralise or to decentralise: An overview of the most recent trends in wastewater treatment management. Journal of Environental Management 94(1) 61-68.

Libralato, G., Volpi Ghirardini, A., Avezzù, F., 2012b. To centralise or to decentralise: An overview of the most recent trends in wastewater treatment management. Journal of Environmental Management 94(1) 61-68.

Lieberherr, E., Klinke, A., Finger, M., 2012. Towards legitimate water governance? The partially privatized Berlin waterworks. Public Management Review 14(7) 923-946.

Lienert, J., Koller, M., Konrad, J., McArdell, C.S., Schuwirth, N., 2011. Multiple-Criteria Decision Analysis Reveals High Stakeholder Preference to Remove Pharmaceuticals from Hospital Wastewater. Environmental Science and Technology (45) 3848-3857.

Lienert, J., Monstadt, J., Truffer, B., 2006. Future scenarios for a sustainable water sector: a case study from Switzerland. Environmental Science & Technology 40(2) 436-442.

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. Additional information for "Structured decision making for sustainable water infrastructure planning under four future scenarios".

Lienert, J., Scholten, L., Egger, C., Maurer, M., 2014b. Structured decision making for sustainable water infrastructure planning under four future scenarios. Under Review.

Lienert, J., Scholten, L., Egger, C., Maurer, M., submitted. Structured decision making for sustainable water infrastructure planning under four future scenarios.

Likert, R., 1932. A technique for the measurement of attitudes. Archives of Psychology(140) 1-55.

Lindhe, A., Rosen, L., Norberg, T., Rostum, J., Pettersson, T.J.R., 2013. Uncertainty modelling in multi-criteria analysis of water safety measures. Environment Systems & Decisions 33(2, Sp. Iss. SI) 195-208.

Linkov, I., Moberg, E., 2012. Multi-Criteria Decision Analysis. Environmental Applications and Case Studies. CRC Press, Taylor & Francis Group, Boca Ratón, FL.

Liu, Z., Kleiner, Y., Rajani, B., Wang, L., Condit, W., 2012. Condition Assessment Technologies for Water Transmission and Distribution Systems. EPA United States Environmental Protection Agency.

Low Choy, S., O'Leary, R., Mengersen, K., 2009. Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology 90(1) 265-277.

Lundie, S., Peters, G.M., Beavis, P.C., 2004. Life cycle assessment for sustainable metropolitan water systems planning. Environmental Science and Technology 38(13) 3465-3473.

Mailhot, A., Pelletier, G., Noel, J.F., Villeneuve, J.P., 2000. Modeling the evolution of the structural state of water pipe networks with brief recorded pipe break histories: Methodology and application. Water Resources Research 36(10) 3053-3062.

Mailhot, A., Poulin, A., Villeneuve, J.P., 2003. Optimal replacement of water pipes. Water Resources Research 39(5).

Makropoulos, C.K., Butler, D., 2010. Distributed Water Infrastructure for Sustainable Communities. Water Resources Management (24) 2795-2816.

Marlow, D.R., Beale, D.J., Burn, S., 2010. A pathway to a more sustainable water sector: Sustainability-based asset management. Water Science and Technology 61(5) 1245-1255.

Martin, P., 2009. Wiederbeschaffungswert der Umweltinfrastruktur. Umfassender Überlick für die Schweiz. (Replacement value of the environmental infrastructure. Overview for Switzerland). Umwelt-Wissen Nr. 0920. Berne, Federal Office for the Environment. http://www.bafu.admin.ch/publikationen/publikation/01058/index.html. Accessed 28 March 2011

Martin, W.E., Bender, H.W., Shields, D.J., 2000. Stakeholder objectives for public lands: Rankings of forest management alternatives. Journal of Environmental Management (58) 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 (42) 1026-1042.

Maurer, M., Rothenberger, D., Larsen, T.A., 2005. Decentralised wastewater treatment technologies from a national perspective: at what cost are they competitive? . Water Science and Technology: Water Supply 5(6) 145-154.

Mays, L.W., 1996. Review of reliability analysis of water distribution systems, Stochastic hydraulics. Proc. symposium: Mackay, Australia, pp. 53-62.

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(230) 1-22.

MeteoSwiss, 2011. IDAweb- The data portal of MeteoSwiss for research and teaching. http://www.meteoschweiz.admin.ch/web/en/services/data_portal/idaweb.html

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(5863) 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(1-3) 5-20.

Morgan, M.G., Henrion, M., 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge.

Morton, A., Fasolo, B., 2009. Behavioural decision theory for multi-criteria decision analysis: a guided tour. Journal of the Operational Research Society 60(2) 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(6).

Mustajoki, J., 2012. Effects of imprecise weighting in hierarchical preference programming. European Journal of Operational Research 218(1) 193-201.

Mustajoki, J., Hämäläinen, R.P., 2000. Web-HIPRE: Global decision support by value tree and AHP analysis. INFOR 38(3) 208-220.

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., Salo, A., 2005. Decision support by interval SMART/SWING - Incorporating imprecision in the SMART and SWING methods. Decision Sciences 36(2) 317-339.

Nash, J.C., Varadhan, R., 2011. Unifying Optimization Algorithms to Aid Software System Users: optimx for R. Journal of Statistical Software 43(9) 1-14.

Nutt, P., 2004. Expanding the search for alternatives during strategic decision making. Academy of Management Executive 18(4) 13-28.

O'Hagan, A., 2012. Probabilistic uncertainty specification: Overview, elaboration techniques and their application to a mechanistic model of carbon flux. Environmental Modelling & Software 36(0) 35-48.

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.

O'Hagan, A., Oakley, J.E., 2004. Probability is perfect, but we can't elicit it perfectly. Reliability Engineering and System Safety 85(1-3) 239-248.

O'Leary, R.A., Choy, S.L., Murray, J.V., Kynn, M., Denham, R., Martin, T.G., Mengersen, K., 2009. Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby Petrogale penicillata. Environmetrics 20(4) 379-398.

Oakley, J.E., O'Hagan, A., 2007. Uncertainty in prior elicitations: a nonparametric approach. Biometrika 94(2) 427-441.

Oakley, J.E., O'Hagan, A., 2010. SHELF: the Sheffield Elicitation Framework (version 2.0): School of Mathematics and Statistics, University of Sheffield.

Osborn, A.F., 1963. Applied imagination: Principles and Procedured of Creative Problem Solving, 3rd ed. C.Scribner, New York.

Palme, U., Lundin, M., Tillman, A.-M., Molander, S., 2005. Sustainable development indicators for wastewater systems – researchers and indicator users in a co-operative case study. Resources Conservation and Recycling (43) 293-311.

Pan, W., Chappell, R., 1998. A Nonparametric Estimator of Survival Functions for Arbitrarily Truncated and Censored Data. Lifetime Data Analysis 4(2) 187-202.

Patz, J.A., Vavrus, S.J., Uejio, C.K., McLellan, S.L., 2008. Climate change and waterborne disease risk in the Great Lakes region of the US. American Journal of Preventive Medicine (35) 451-458.

Pelletier, G. (2000) Impact du remplacement des conduited d'aqueduc sur le nombre annuel de bris (PhD thesis), Université du Québec, Québec.

Pelletier, G., Mailhot, A., Villeneuve, J.P., 2003. Modeling water pipe breaks - three case studies. Journal of Water Resources Planning and Management 129(2) 115-123.

Pennings, J.M.E., Garcia, P., 2009. The informational content of the shape of utility functions: Financial strategic behavior. Managerial and Decision Economics 30(2) 83-90.

Peterson, G.D., Cumming, G.S., Carpenter, S.R., 2003. Scenario planning: A tool for conservation in an uncertain world. Cons Biol 17(2) 358-366.

Poulton, M., Le Gat, Y., Brémond, B., 2007. The impact of pipe segment length on break predictions in water distribution systems, In: Alegre, H., do Céu Almeida, M. (Eds.), Strategic Asset Management of Water Supply and Wastewater Infrastructure. IWA Publishing.

Pöyhönen, M., Vrolijk, H., Hämäläinen, R.P., 2001. Behavioral and procedural consequences of structural variation in value trees. European Journal of Operational Research 134(1) 216-227.

Pujol, G., Iooss, B., Janon, A. (2012) Sensitivity: Sensitivity Analysis. R package version 1.6-1. http://CRAN.R-project.org/package=sensitivity.

Qualtrics (2012) Qualtrics Research Suite. 2.227s. http://www.qualtrics.com/.

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

Rabin, M., 2000. Risk Aversion and Expected-Utility Theory: A Calibration Theorem. Econometrica 68(5) 1281-1292.

Rajani, B., Kleiner, Y., 2001. Comprehensive review of structural deterioration of water mains: physically based models. Urban Water 3 151-164.

Raju, K.S., Pillai, C.R.S., 1999. Multicriterion decision making in performance evaluation of an irrigation system. European Journal of Operational Research (112) 479-488.

Raju, K.S., Vasan, A., 2007. Multi attribute utility theory for irrigation system evaluation. Water Resources Management 21(4) 717-728.

Ram, C., Montibeller, G., 2012. Exploring the impact of evaluating strategic options in a scenario-based multi-criteria framework. Technological Forecasting and Social Change 80(4) 657-672.

Ram, C., Montibeller, G., 2013. Exploring the impact of evaluating strategic options in a scenario-based multi-criteria framework. Technological Forecasting and Social Change 80(4) 657-672.

Ram, C., Montibeller, G., Morton, A., 2011. Extending the use of scenario planning and MCDA for the evaluation of strategic options. Journal of the Operational Research Society 62(5) 817-829.

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 and Software 22(11) 1543-1556.

Rehan, R., Knight, M.A., Haas, C.T., Unger, A.J.A., 2011. Application of system dynamics for developing financially self-sustaining management policies for water and wastewater systems. Water Research 45(16) 4737-4750.

Reichert, P., Borsuk, M.E., Hostmann, M., Schweizer, S., Spörri, C., Tockner, K., Truffer, B., 2007. Concepts of decision support for river rehabilitation. Environmental Modelling & Software(22) 188-201.

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.

Renaud, E., De Massiac, J.C., Brémond, B., Laplaud, C., 2009. SIROCO, a decision support system for rehabilitation adapted for small and medium size water distribution companies, In: Alegre, H., do Céu Almeida, M. (Eds.), Strategic Asset Management of Water Supply and Wastewater Infrastructures. IWA Publishing: London, pp. 329-344.

Renaud, E., Le Gat, Y., Poulton, M., 2012. Using a break prediction model for drinking water networks asset management: From research to practice. Water Science and Technology: Water Supply 12(5) 674-682.

Renner, R., Schneider, F., Hohenwallner, D., Kopeining, C., Kruse, S., Lienert, J., Link, S., Muhar, S., 2013. Challenges and coping strategies for transdisciplinary knowledge production towards sustainable water governance. submitted.

Rinderknecht, S.L., Borsuk, M.E., Reichert, P., 2011. Eliciting density ratio classes. International Journal of Approximate Reasoning 52(6) 792-804.

Rinderknecht, S.L., Borsuk, M.E., Reichert, P., 2012. Bridging uncertain and ambiguous knowledge with imprecise probabilities. Environmental Modelling & Software 36(0) 122-130.

Ringland, G., 2002. Scenario Planning Managing for the Future. John Wiley & Sons, Chichester.

Roscher, H., Ahrens, J., Rödiger, S., Sorge, H.-C., 2008. Sanierung und Erneuerung von Wasserrohrnetzen, In: Weimar, B.-U. (Ed.), Rehabilitation von Rohrleitungen- Sanierung und Erneuerung von Ver- und Entsorgungsnetzen, pp. 63-204.

Roscher, H., Mischner, J., Sorge, H.-C., 2005. Vergleich von Bewertungssystemen für Rohrleitungen: Wasser- und Gasrohrnetze. (Comparison of piping network evaluations systems: water and gas piping networks). Erfurt,

Rothschild, M., Stiglitz, J.E., 1970. Increasing risk: I. A definition. Journal of Economic Theory 2(3) 225-243.

Roy, B., 1991. The outranking approach and the foundations of electre methods. Theory and Decision 31(1) 49-73.

Roy, B., 2010. Robustness in operational research and decision aiding: A multi-faceted issue. European Journal of Operational Research 200(3) 629-638.

Runge, M.C., Converse, S.J., Lyons, J.E., 2011. Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. Biological Conservation 144(4) 1214-1223.

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 (21) 1031-1045.

Sa-nguanduan, N., Nititvattananon, V., 2010. Strategic decision making for urban water reuse application: A case from Thailand. Desalination 268(1-3) 141-149.

Sægrov, S., 2005. CARE-W: computer aided rehabilitation of water networks. IWA Publishing, London.

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

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(2) 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(10-11) 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(3) 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(1) 39-56.

Scheidegger, A. (2012) adaptMCMC: Implementation of a generic adaptive Monte Carlo Markov Chain sampler. v1.0.3. CRAN R project. http://CRAN.R-project.org/package=adaptMCMC.

Scheidegger, A., Hug, T., Rieckermann, J., Maurer, M., 2011. Network condition simulator for benchmarking sewer deterioration models. Water Research 45(16) 4983-4994.

Scheidegger, A., Maurer, M., 2012. Identifying biases in deterioration models using synthetic sewer data. Water Science and Technology 66(11) 2363-2369.

Scheidegger, A., Scholten, L., Maurer, M., Reichert, P., 2013. Extension of pipe failure models to consider the absence of data from replaced pipes. Water Research 47(11) 3696-3705.

Schmidt, U., Starmer, C., Sugden, R., 2008. Third-generation prospect theory. Journal of Risk and Uncertainty 36(3) 203-223.

Schnaars, S.P., 1987. How to develop and use scenarios. Long Range Planning 20(1) 105-114.

Schoemaker, P.J.H., 1995. Scenario Planning: A Tool for Strategic Thinking. Sloan Management Review 36(2) 25-40.

Scholten, L., Reichert, P., Schuwirth, N., Lienert, J., submitted. Tackling uncertainty in multi-criteria decision analysis- An application to water supply infrastructure planning.

Scholten, L., Scheidegger, A., Reichert, P., Maurer, M., 2013. Combining expert knowledge and local data for improved service life modeling of water supply networks. Environmental Modelling & Software 42 1-16.

Scholten, L., Scheidegger, A., Reichert, P., Maurer, M., Lienert, J., 2014. Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis. Water Research 49 124-143.

Schulz, M., Renn, O., 2009. Das Gruppendelphi; Konzept und Fragebogenkonstruktion. VS Verlag für Sozialwissenschaften, GWV Fachverlage GmbH: Wiesbaden, Germany.

Schuwirth, N., Reichert, P., Lienert, J., 2012a. 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(2) 472-483.

Schuwirth, N., Stamm, C., Reichert, P., 2012b. Incorporation of uncertainty in decision support to improve water quality, In: Seppelt, R., Voinov, A.A., Lange, S., Bankamp, D. (Eds.), International Congress on Environmental Modelling and Software: Sixth Biennial Meeting, Leipzig, Germany.

Selvakumar, A., Tafuri, A.N., 2012. Rehabilitation of Aging Water Infrastructure Systems: Key Challenges and Issues. Journal of Infrastructure Systems 18(3) 202-209.

Shalit, H., Yitzhaki, S., 1994. Marginal Conditional Stochastic Dominance. Management Science 40(5) 670-684.

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(4) 577-583.

Smidts, A., 1997. The relationship between risk attitude and strength of preference: A test of intrinsic risk attitude. Management Science 43(3) 357-370.

Speirs-Bridge, A., Fidler, F., McBride, M., Flander, L., Cumming, G., Burgman, M., 2010. Reducing Overconfidence in the Interval Judgements of Experts. Risk Analysis 30(3) 512-523.

Spetzler, C.S., Stael von Holstein, C.-A.S., 1975. Probability Encoding in Decision Analysis. Management Science 22(3) 340-358.

Stadt Zürich, 2012. Jahresbericht 2011: Seewasser-Untersuchungen. http://www.stadt-zuerich.ch/dib/de/index/wasserversorgung/wasserqualitaet/qualitaetswerte.html. Accessed 20 September 2012

Sterman, J.D., 2000. Business dynamics: systems thinking and modeling for a complex world. McGraw-Hill, Boston.

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

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

Stewart, T.J., French, S., Rios, J., 2013. Integrating multicriteria decision analysis and scenario planning—Review and extension. Omega 41(4) 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 (76) 1150-1162.

SVGW, 2009. Statistische Erhebungen der Wasserversorgungen in der Schweiz Betriebsjahr 2008. Zürich, Schweizer Verein des Gas- und Wasserfaches.

Swart, R.J., Raskin, P., Robinson, J., 2004. The problem of the future: Sustainability science and scenario analysis. Global Environ Change – Human Policy Dimens 14(2) 137-146.

SWIP, 2013. Sustainable Water Infrastructure Planning. Project of the National Research Program NRP61. http://www.eawag.ch/forschung/sww/schwerpunkte/infrastrukturen/planung_wasserinfrastr/index_EN. Accessed 02 December 2013

Swissplan, 2012. Finanzmanagement in der Siedlungswasserwirtschaft - Kurzbericht zum Normalhaushalt 2010. (Financial management in urban water management). http://www.awel.zh.ch/internet/baudirektion/awel/de/wasserwirtschaft/wasserversorgung/finanzierung/zahlen.html#a-content. Accessed 4 August 2013

Ten Veldhuis, J., Clemens, F., Sterk, G., Berends, B., 2010. Microbial risk associated with exposure to pathogens in contaminated urban flood water. Water Research 44(9) 2910-2918.

Thomson, J., Wang, L., 2009. Condition Assessment of ferrous water transmission and distribution systems. State of technology review report. Office of Research and Development, National Risk Management Research Laboratory.

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(7) 702-722.

TREMOD, 2010. Fortschreibung und Erweiterung Daten- und Rechenmodell. Energieverbrauch und Schadstoffemissionen des motorisierten Verkehrs in Deutschland 1960 - 2030 (TREMOD, Version 5). http://www.probas.umweltbundesamt.de/php/themen.php?&prozessid={6D34A996-4DCC-7061-9B52-3006450AB34A}&id=13455327232&step=4&search=. Accessed 11 July 2013

Truffer, B., Störmer, E., Maurer, M., Ruef, A., 2010. Local strategic planning processes and sustainability transitions in infrastructure sectors. Environmental Policy and Governance 20(4) 258-269.

Trujillo Alvarez, R., 1995. Bedarfsprognose und Strategieentwicklung für die Rehabilitation städtischer Wasserrohrnetze. Schriftenreihe des Instituts für Städtebau und Landesplanung, Universität Fridericiana zu Karlsruhe.
Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Biases in judgments reveal some heuristics of thinking under uncertainty. Science 185(4157) 1124-1131.
Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211(4481) 453-458.
U.S. Government Accountability Office, 2004. Water infrastructure. Comprehensive Asset Management Has Potential to Help Utilities Better Identify Needs and Plan Future Investments. Publication No. GAO-04-461. http://www.gpoaccess.gov/gaoreports/index.html. Accessed 22 May 2013
Ugarelli, R., Venkatesh, G., Brattebø, H., Di Federico, V., Sægrov, S., 2010. Asset management for urban wastewater pipeline networks. Journal of Infrastructure Systems 16(2) 112-121.
Vanier, D., 2001. Why Industry Needs Asset Management Tools. Journal of Computing in Civil Engineering 15 (Special Issue: information technology for life-cycle infrastructure management) 35-43.
Vihola, M., 2011. Robust adaptive Metropolis algorithm with coerced acceptance rate. Statistics and Computing 1-12.
Vince, F., Aoustin, E., Breant, P., Marechal, F., 2008. LCA tool for the environmental evaluation of potable water production. Desalination 220(1-3) 37-56.
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.
Vreeburg, J.H.G., Blokker, E.J.M., Horst, P., Van Dijk, J.C., 2009. Velocity-based self-cleaning residential drinking water distribution systems. Water Science and Technology: Water Supply 9(6) 635-641.
VSA, 2002. Regenwasserentsorgung. Richtlinie zur Versickerung, Retention und Ableitung von Niederschlagswasser in Siedlungsgebieten. (Rain water disposal. Guideline for infiltration, retention and discharge of stormwater in urban areas). Verband Schweizerischer Abwasser- und Gewässerschutzfachleute (Association of Swiss wastewater and water protection experts). http://www.vsa.ch/en/publikationen/shop/. Accessed 09 January 2013
VSA, 2007. Erhaltung von Kanalisationen. Richtlinien Merkblätter und Empfehlungen.
VSA, 2011. Kosten und Leistungen der Abwasserentsorgung. (Costs and services of wastewater disposal). Association of Swiss wastewater and water protection experts. http://www.vsa.ch/en/publikationen/shop. Accessed 19 December 2012
Wack, P., 1985. Scenarios: Uncharted water ahead. Harvard Business Review 63(5) 72-89.
Wakker, P., Deneffe, D., 1996. Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown. Management Science 42(8) 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(1) 5 - 17.
Walker, W.E., Marchau, V.A.W.J., Swanson, D., 2010. Addressing deep uncertainty using adaptive policies: Introduction to section 2. Technological Forecasting and Social Change 77(6) 917-923.
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(7) 1336-1349.
Watson, T.G., Christian, C.D., Mason, A.J., Smith, M.H., Meyer, R., 2004. Bayesian-based pipe failure model. Journal of Hydroinformatics 6(4) 259-264.
Wayne, N.B., 2004. Applied Life Data Analysis. John Wiley & Sons Inc.
WCED, 1987. World Commission on Environment and Development. Our Common Future. Oxford University Press, Oxford.
Weber, M., 1987. Decision making with incomplete information. European Journal of Operational Research(28) 44-57.
Weber, M., Borcherding, K., 1993. Behavioral influences on weight judgments in multiattribute decision-making. European Journal of Operational Research 67(1) 1-12.
Wickham, H., 2007. Reshaping data with the reshape package. Journal of Statistical Software 21(12).
Wickham, H., 2009. ggplot2: elegant graphics for data analysis. Springer New York.

Winkler, R.L., 1967. The assessment of prior distributions in Bayesian analysis. Journal of the American Statistical Association 62(319) 776-880.

Wong, T.H.F., Brown, R.R., 2009. The water sensitive city: Principles for practice. Water Science and Technology 60(3) 673-682.

Wuelser, G., Pohl, C., Hirsch, H.G., 2012. Structuring complexity for tailoring research contributions to sustainable development: a framework. Sustainability Science 7(1) 81-93.

Yitzhaki, S., 2003. Gini's Mean difference: a superior measure of variability for non-normal distributions. METRON- International Journal of Statistics LXI(2) 285-316.

Zheng, J., Egger, C., Lienert, J., 2013. Incorporating stakeholders' preferences in wastewater infrastructure planning in Switzerland (working title). In preparation.

Zhou, P., Ang, B.W., 2009. Comparing MCDA Aggregation Methods in Constructing Composite Indicators Using the Shannon-Spearman Measure. Social Indicators Research 94(1) 83-96.

SI-A) Additional Information for "Structured decision making for sustainable water infrastructure planning under four future scenarios"

Abstract

To support Sustainable Water Infrastructure Planning (SWIP), a participatory decision-making procedure was developed in the SWIP-project at Eawag[1]. This procedure is based on Structured Decision Making (SDM)[2], which guides stakeholders through different steps of the decision process: (1) clarify decision context (may include a stakeholder analysis); (2) define objectives and attributes; (3) develop alternatives; (4) estimate consequences; (5) evaluate trade-offs and select alternatives; and (6) implement, monitor and review.

The SDM-application to water infrastructure planning was developed in close collaboration with stakeholders in a case study near Zürich, Switzerland. The experienced advantages and disadvantages of the first steps of the proposed procedure were discussed in a scientific publication by Lienert et al. (2014)[3]. We strongly encourage others to apply this SDM-procedure for sustainable water infrastructure planning to their specific case. To this end, the approach was developed in a generalized way and we present more material covering the different steps of the SDM-procedure in this Working Paper.

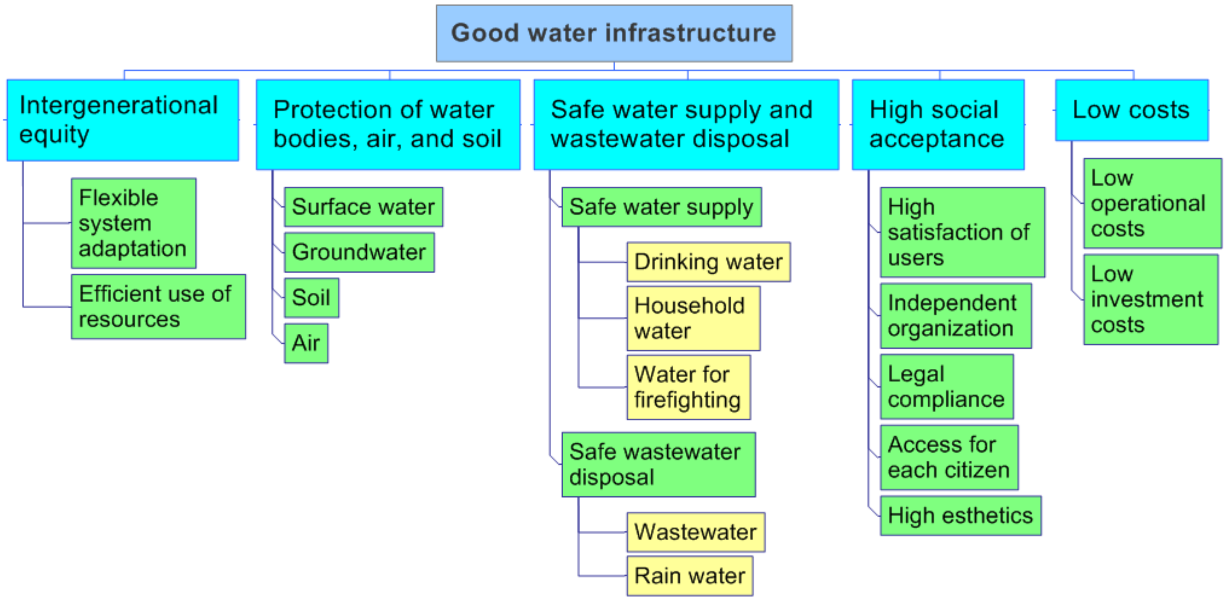
The here presented material includes different steps in the development of a comprehensive objectives hierarchy for water supply and wastewater management. The objectives are operationalized with attributes (indicators/ benchmarks), which are described in detail, including the ranges (best- and worst-possible case) and a description of the Status Quo. Four future scenarios were developed in a scenario planning workshop together with local stakeholders to capture socio-demographic uncertainty, which are again described in detail. Ten strategic decision alternatives were developed by stakeholders with help of a strategy generation table. These include the current system with central water supply and wastewater treatment plants, but also fully decentralized on-site options and different management strategies. The strategy generation table can be used to tailor decision alternatives for water infrastructure planning to other cases. Finally, we provide detailed feedback from the stakeholders for each step. We evaluate the proposed SDM-approach and give recommendations for other applications.

Keywords: decision making, water infrastructure, scenario planning, stakeholder participation, structuring, water management

A1 Step (2) Define objectives and attributes

1. Preliminary objectives hierarchy

Preliminary objectives hierarchy created by the project team and discussed with the stakeholders in the 27 face-to-face interviews. Details concerning this interview series and the stakeholder selection are given in the stakeholder and network analysis (Lienert et al., 2013).



2. Face-to-face interviews concerning objectives

In the face-to-face interviews with 27 stakeholders (see Lienert et al. 2014), each stakeholder was asked to classify the objectives into: essential (without this objective I cannot judge whether the fundamental objective is reached), important (without this it would be difficult to judge whether the fundamental objective was reached), and nice to have (attainment of fundamental objective can be judged without). NS = not significant for this stakeholder or missing (e.g. the water supply objectives were not judged by the wastewater stakeholders). The water supply and wastewater objectives were split and only judged by the respective stakeholders (explaining the large number of NS in Tabs. 1 and 2). The results concerning the objectives on the highest-level of the hierarchy are given in Table 1 and for the lower-level fundamental objectives in Table 2.

The objective "low costs" was judged as only "nice to have" by ten interviewees. However, the corresponding two lower-level objectives ("low operational" and "low investment costs") were judged as "nice to have" by only three and four stakeholders, respectively, which seems a bit contradictory. Most of the other 18 lower-level objectives were judged as very important by the large majority, with exception of "protection of air", which was classified as only "nice to have" by seven interviewees.

Based on the input from the interviews and an extensive discussion in the scientific project team, the objectives hierarchy was again revised. The revised version is given in Figure 2. This version was used as input for the stakeholder workshop.

Table 29: Classification of highest level objectives in interviews. We show the results of face-to-face interviews with 27 stakeholders about the importance of the highest-level fundamental objectives. Details see text.

Objective → Classification ↓	Intergenerational equity	Protection of water, air and soil	Safe water supply	Safe wastewater disposal	High social acceptance	Low costs
Essential	4	17	18	13	4	3
Important	16	7	1	1	17	10
Nice to have	2	0	0	0	3	10
NS or missing	5	3	8	13	3	4

Table 30: Classification of the lower-level objectives in interviews, see Table 27.

Objective	**Intergenerational equity**		**Protection of water, air and soil**				**Safe water supply**			**Safe wastewater disposal**		**High social acceptance**					**Low costs**	
Sub-objective → Classification ↓	Flexible system adaptation	Efficient use of resources	Surface water	Ground-water	Soil	Air	Drinking water	House-hold water	Water for firefighting	Waste-water	Rain water	High satisfaction of users	Indepen-dent or-ganization	Legal com-pliance	Access for each citizen	High esthetics	Low operation-al costs	Low investment costs
Essential	7	8	14	17	7	4	18	7	9	13	9	10	1	12	11	4	15	8
Important	15	14	7	4	6	6	1	8	8	1	5	10	4	3	8	8	6	11
Nice to have	1	1	1	1	3	7	0	1	0	0	0	4	5	4	1	4	3	4
NS or missing	4	4	5	5	11	10	8	11	10	13	13	3	17	8	7	11	3	4

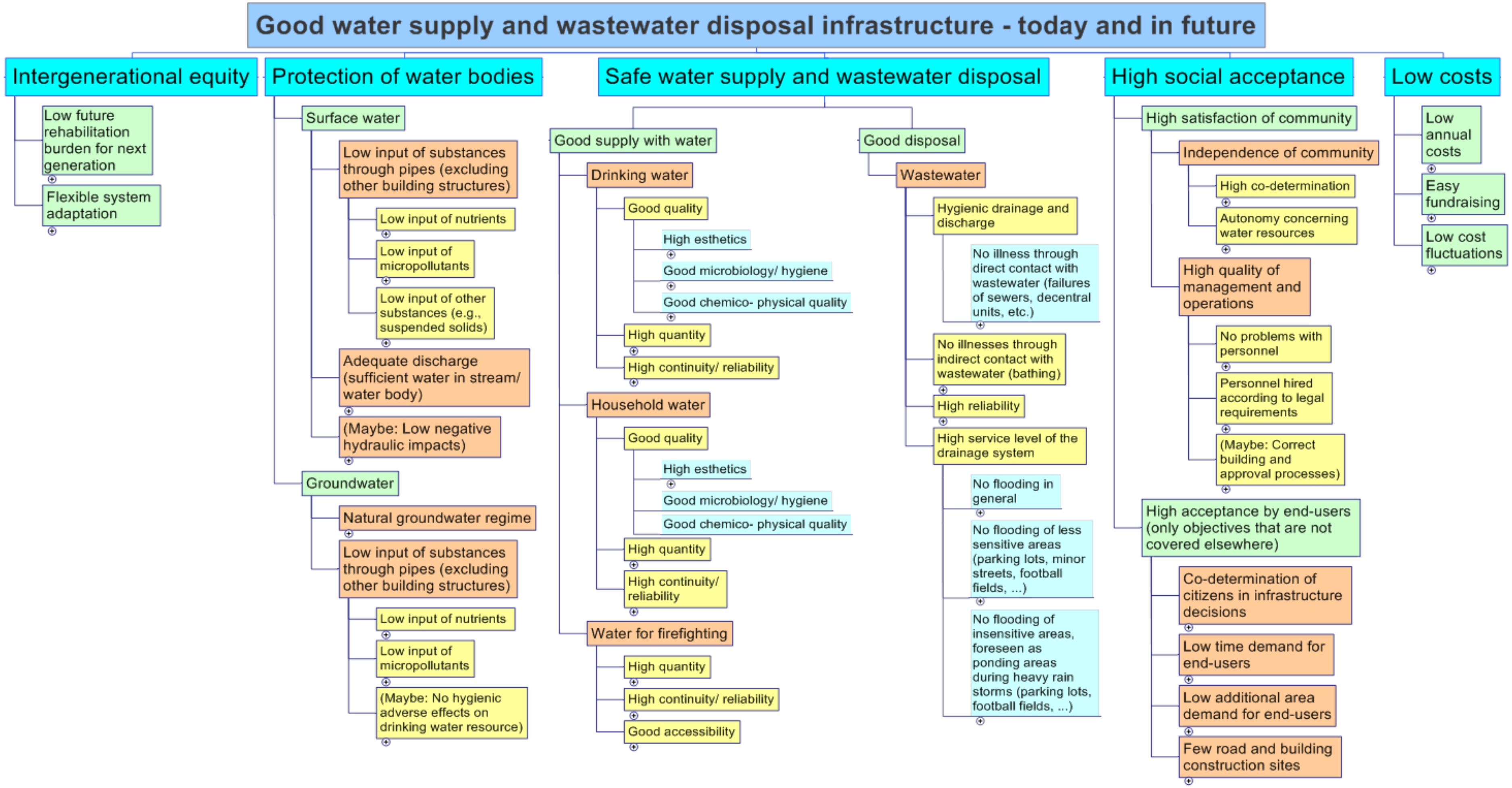


Figure 25: Objectives hierarchy using in workshop, after revision by the project team. Revision based on input from interviews (Tab. 26 and 27) and extensive discussions in the scientific project team. The highest-level (1) fundamental objectives are colored blue, the second-level (2) are green, the third-level (3) are orange, the fourth-level (4) are yellow and the lowest-level fundamental objectives (5) are colored light blue.

3. Discussion of objectives in workshop

The objectives hierarchy presented in Figure 24 was extensively discussed in a stakeholder workshop (see Lienert et al. 2014). In Table 28, we show the feedback and group discussions concerning the proposed fundamental objectives.

Table 31: Feedback in workshop concerning fundamental objectives. We show objectives on different levels of an objectives hierarchy (see Fig. 2), where level 1 corresponds to the highest level (in blue, e.g. Intergenerational equity, Protection of water bodies,…), level two to the next-lower level (in green, e.g. Low rehabilitation burden, Flexible system adaptation,…) etc. Furthermore, we present the proposed attributes, the written feedback by workshop participants as discussed in groups of two, the main issues discussed in the plenary in the workshop, and the distribution of points (each of 20 workshop participants received three points which he could distribute to mark those objectives that seemed least relevant). Propositions for new objectives are given at the end of the table. WWTP = wastewater treatment plant; CSO = combined sewer overflows.

	Objective (level 1)	Objective (level 2)	Objective (level 3)	Attribute	Written notes from workshop participants	Discussion in workshop	Points
1	Intergenerational equity				One comment that this objective is OK and that long-term investments should be made according to today's already existing concepts; one comment that it can be deleted because it is already contained in "safe water supply and wastewater disposal"		
	Low future rehabilitation burden for next generation			Is rehabilitation demand during this generation done in this generation? (e g % necessary realization)	One comment that this is not a relevant objective		1
	Flexible system adaptation			Ease of technical extension or deconstruction of infrastructure (expert predictions)	Four comments: flexible adaptation can be deleted; additional comments that the system is very inert and changes are slow; flexibility is expensive; the uncertainties remain; is dependent on technological innovations, which are not foreseeable		6
2	Protection of water bodies				"Protection of soil and air" was removed before the workshop by project team; they were considered less important by many interview partners and also the project team		
	Surface water	Low input of substances through pipes (excluding other building structures)	Low input of nutrients	g/m³; kg/a	Some ("fantastic") alternatives might not have pipes anymore; hence this is not a good objective; several statements that the objective should be less specific, just state that there should be a reduction of the pollution from wastewater	Discussion that other building structures such as infiltration structures and WWTP should be included since they are relevant for input of substances; discussion about the system boundaries of analysis?	
			Low input of micropollutants	µg/m³; g/a			1
			Low input of other substances (e g suspended solids)	µg/m³; g/a	Two statements that "other substances" are not important, can be deleted		

Objective (level 1)	Objective (level 2)	Objective (level 3)	Attribute	Written notes from workshop participants	Discussion in workshop	Points
	Adequate discharge (sufficient water in stream/ water body)		L/s; /(s*h)	Two statements that discharge is not relevant because there is only discharge into river during rain events when there is sufficient water in the river / stream anyway	Controversy between: discharge is not so relevant because of natural variations of river discharge and rainfall and: discharge is relevant because there can be too low dilution of substances from urban areas	1
	(Maybe: Low negative hydraulic impacts)		Number of bed-moving floods due to CSOs / no of bed-moving floods without CSOs	Four statements that it can be deleted because natural bed-moving floods are much larger than those from CSOs; two suggestions to change it to "no mechanical negative impacts"	Not so relevant because there are natural variations in discharge due to rainfall	3
Groundwater	Natural groundwater regime		% Removal / regeneration	One statement that groundwater is not relevant; one that this is critical, but difficult / impossible to measure		
	Low input of substances through pipes (excluding other building structures)	Low input of nutrients	g/m³; kg/a	One comment that "low input of substances" is sufficient, without distinguishing between nutrients and micropollutants; OK to not include details of WWTP; relevant objective if there is a dynamic development of the settlement		
		Low input of micropollutants	µg/m³; g/a			
	(Maybe: No hygienic adverse effects on drinking water resource)		Semi-quantitative expert estimate (state of pipes; prob that pumps break, etc)	Two comments that hygienic effects on groundwater can be deleted; one comment that drinking water protection zones are relevant and should be included		
3 Safe water supply (good supply with water)						
Drinking water	Good quality	High esthetics	Taste, smell, etc	Can be subjective, e g in USA chloride characterizes safe water		
		Good microbiology / hygiene	*Escherichia coli*, *Pseudomonas aeruginosa*, enterococci in colony forming units (CFU/100ml)			
			Potential of re-contamination			
		Good chemico-physical quality	Inorganic substances (N-compounds)			
			Turbidity			
			Pesticides, micropollutants			
			Dinitrophenols			
			Corrosion potential of metals			

Objective (level 1)	Objective (level 2)	Objective (level 3)	Attribute	Written notes from workshop participants	Discussion in workshop	Points
	High quantity		L/(person*d)	Comment that it can be deleted/ that it is relevant and that the quantity should be multiplied by 3 to include water for industry		
	High continuity / reliability		Customer minutes lost (length of outage * number of people affected/1000 people)			
			Hours with outage			
Household water						4
	Good quality	High esthetics	Taste, smell, etc	Three comments that high esthetics of household water can be deleted, because it is not relevant, e g for the washing machine		2
		Good microbiology / hygiene	*Escherichia coli*, *Pseudomonas aeruginosa*, enterococci in CFU/ 100ml	Two comments that good microbiology / hygiene is irrelevant for household water		1
			Potential of re-contamination			
		Good chemico-physical quality	Anorganic substances (N-compounds)			
			Turbidity			
			Pesticides, micropollutants			
			Dinitrophenols			
			Corrosion potential of metals			
	High quantity		L/(person*d)	Could be changed to "sufficient quantity"		
	High continuity/ reliability		Customer minutes lost (length of outage * number of people affected/1000 people)			
			Hours with outage			
Water for firefighting						3
	High quantity		l/min with minimally 3,5 bar flow pressure	One comment that "water for firefighting" can be deleted	High pressure of water for firefighting is important	1
			Water reserve m^3 per pressure zone			
			Flow rate l/s			

Objective (level 1)	Objective (level 2)	Objective (level 3)	Attribute	Written notes from workshop participants	Discussion in workshop	Points
	High continuity / reliability		Criticality index			
	Good accessibility		Length of hose to building			
4 Safe wastewater disposal (good disposal)						
Wastewater	Hygienic drainage and discharge	No illness through direct contact with w-water (failures sewers, decentral units, etc)	e g Number of illnesses in population per year	Two comments that double coverage with number of illnesses can be deleted	Maintenance-friendliness should be included (e g easy access to manholes, easy to flush, …)	
		No illnesses through indirect contact with wastewater (bathing)	e g Number of illnesses in population per year			
	High reliability		Customer minutes lost (Length of outage * number of people affected)			
	High service level of the drainage system	No back pressure of wastewater (anywhere)	Number of people affected * length of back pressure	One comment: "no back pressure" is unrealistic; several comments: all 3 sub-obj needed for non-conventional solutions	Damages should be included	
		No back pressure of rain on retention areas (parking lots, football fields,)	Number of people affected * length of back pressure	Suggestion to change to "few" or "controlled" back pressures; four comments that it can be deleted		4
		No uncontrolled back pressure of rain (e g streets, non-retention parking lots)	Number of people affected * length of back pressure	Two comments that this can be deleted		
5 High social acceptance					Social acceptance is not so important; is a soft factor Can be assessed for today, but difficult for the future	1
High satisfaction of community						7
	Independence of community	High co-determination	Qualitative: influence of community (differs between organizational forms)	One comment that co-determination of community is too political and should be deleted; one comment that direction is unclear (is more or less better?)	Acceptance by community depends on people working there; independence was so far not important in infra-structure decisions Currently, no co-determination for telecommunication, but it works well Discussion whether it can be deleted; but objective is necessary to measure organizational forms of some alternatives	
		Autonomy concerning water resources	% Of annual water demand from external providers	Eight comments that autonomy of water resource is not important and can be deleted		3
	High quality of	No problems with	Number of working hours	Two comments that this can be deleted; one comment that	Is strongly dependent on personnel, and not on size	3

Objective (level 1)	Objective (level 2)	Objective (level 3)	Attribute	Written notes from workshop participants	Discussion in workshop	Points
	management and operations	personnel	per year required from volunteers	well-trained personnel is important	of the network or professionalism of organization; efficiency is not only measurable in costs Discussion that it is required to distinguish organizational forms	
		Personnel hired according to legal requirements	Number of hours / year that surpass the legally allowed maximal working hours (for stand-by emergency duties)	Two comments that objective is not important; legal requirements should be fulfilled; flexibility of job market is not relevant		2
		(Maybe: Correct building and approval processes)	% Approvals granted for "correct approval process" / total approvals	Eight comments that this objective can be deleted; legal requirements are boundary condition; instead use objective "simplified processes"		1
High acceptance by end-users	Co-determination of citizens in infrastructure decisions		Degree of co-determination (expert estimate; classes)	Three comments that co-determination of citizens is irrelevant	In long term (25 - 40 years), acceptance by end-users is more important than acceptance by community	1
	Low time demand for end-users		hrs /yr			
	Low additional area demand for end-users		Additional area demand on private property per end user (m² or maybe m³ in buildings)	Six comments that this objective can be deleted; it is unimportant, because 98 % is below the ground anyway; public interest is more important than personal interests		3
	Few road and building construction sites		Number of building sites in community / year weighted with average number? or length of pipes?	Three comments that this objective can be deleted		7
6 Low costs				Several comments that the overall annual costs are important; not the details		1
Low annual costs			Capital costs; CHF / year (interest rates, depreciation, investment costs)	General comment that details concerning costs are not important; it is dynamic over the decades; the overall costs are important		
			Personnel costs; CHF / year	One comment that personnel costs can be deleted		
			Material costs; CHF / year	Two comments: "operational costs" are more important than material costs		
Easy fundraising			Qualitative in classes (dependent on size of organization and level of	One comment that it is important that also small organizations receive subventions; general comment that it is not important		3

Objective (level 1)	Objective (level 2)	Objective (level 3)	Attribute	Written notes from workshop participants	Discussion in workshop	Points
			debts)			
Low cost fluctuations			Number of increases >5 % (compared with previous year over 40 years)	Several comments that increase of costs is not so important; only the overall annual costs are important		5
Propositions for other objectives (based on individual written statements in workshop)						
Good supply with drinking water	High water pressure			Mentioned twice		
Safe wastewater disposal	High water pressure			Mentioned twice		
High satisfaction of community	Ease of maintenance			Mentioned twice		
High satisfaction of community	Low additional area demand for community			One comment that this is important, because there can be resistance in the community		
High satisfaction of community	High quality of managem and operations	Highly qualified, well-trained personnel		Mentioned once		
High satisfaction of community	High quality of managem and operations	Separate politics from operation		Mentioned once		

4. Modification of objectives hierarchy

In the months following the discussion of the objectives hierarchy in the stakeholder workshop, the SWIP project team carefully went through all objectives and attributes again. The resulting objectives hierarchy (Fig. 1 main text) is less than the one presented in Fig. 24. This is a result of our efforts to cut down the number of objectives to those that are absolutely essential to characterize the water infrastructure system. We deleted objectives that are of minor importance (for water supply and wastewater infrastructures), which do not help to discriminate between the strategic decision alternatives, or for which it seemed impossible to generate reasonable predictions (neither could we model or estimate them ourselves, nor could we find experts that were capable of giving estimates). If possible, we used other attributes instead. The major changes are given below (minor changes, e.g. concerning the wording are not listed).

- Protection of water bodies / surface water / low input of substances through pipes / low input of nutrients / … micropollutants / … other substances → changed to: Protection of water and other resources/surface water / <u>good chemical state of the watercourse</u>

  Reason: The chemical state of the water bodies is the fundamental objective, while the input of e.g. nutrients is only a means objective. As attribute we use an aggregated measure over a number of indicators (several nutrients and pesticides) in five quality classes. We base our assessment on the procedure developed in the related NRP 61 project iWaQa (Schuwirth, 2012; iWaQa 2013), which in turn draws on existing assessment procedures by water authorities in Switzerland and Germany (see Schuwirth et al., 2012b for references). Because it is difficult to elicit preferences from lay people for attributes that characterize a good chemical status of the river, expert valuations of these single indicators are used for our predictions. The valuation scheme is based on the modeled contribution of chemicals from the wastewater infrastructure system to natural water bodies. As reference points we use existing measurement stations of (AWEL, 2006) with additional reference points added to the model used in (iWaQa, 2013), basically upstream and downstream of urban areas. We also rely on the iWaQa experts for the aggregation and weighting procedure of these attributes to come to an overall description of the chemical state of the watercourse in one of five classes (very bad – to very good; also see Langhans et al., 2013). However, we do ask all our respondents for trade-offs between this and other objectives (i.e. for the scaling constant or weight of this objective).

- Protection of water bodies / surface water / <u>adequate discharge</u> (sufficient water in stream / water body) → deleted

  Reason: Removed after extensive discussion, also with the related project (iWaQa, 2013). Our project SWIP relies strongly on iWaQa to model and quantify the effects of the urban infrastructure system on surface waters. However, there are no clearly defined criteria (attributes) available to assess this objectives' degree of fulfillment. Since we are not able to quantify different outcomes of this objective for the different decision alternatives based on our own models, nor is it being modeled in iWaQa, we decided to delete it.

- Protection of water bodies / surface water / <u>low negative hydraulic impacts</u> → included

  Reason: We had first considered excluding this objective, because the water quality experts from iWaQa (2013) are not using it due to the problem of not being able to translate hydraulic events into negative effects for the water ecosystems. However, there are existing guidelines for wastewater engineers in Switzerland (VSA, 2002), which quantify the ratio of bed load movements with or without storm water discharge. We decided to use these. Our attribute is thus very simple: the % reference points in the river network (of the case study catchment) that fulfill the (VSA, 2002) guidelines for storm water handling. The same reference points as for "good chemical state of watercourse" are used. The Status Quo levels are elicited together with engineering experts.

- Protection of water bodies / groundwater / low input of substances through pipes → replaced with low contamination from sewers

  Reason: Leaky sewers are potential inputs of pollutants into the soil and eventually the groundwater. While this is certainly an important objective, it is very difficult to quantify, since it is dependent on various factors. We decided to use semi-quantitative expert judgments (groundwater specialists at Eawag) to estimate the amount of wastewater exfiltrating from sewer lines, dependent on their physical condition. The assessment of the attribute in terms of water quality classes follow those used for a "good chemical state of watercourses".

- Protection of water bodies / groundwater / → additional objective low contamination from infiltration structures

  Reason: Additionally to leaky sewers, infiltration of storm water from impervious areas such as roofs increases the risk of contaminating the groundwater. This risk depends, for example, on the location of the infiltration structure and the amount of rain water being infiltrated. As above, the potential for contamination is based on estimates from groundwater experts in five water quality classes.

- Protection of water bodies → changed to protection of water and other resources; i.e. including two new objectives: Recovery of nutrients and efficient use of electrical energy

  Reason: We decided that these are two important fundamental objectives that should be included in a comprehensive objectives hierarchy, which also focuses on ecological sustainability. Moreover, the recovery of nutrients from wastewater (characterized by the indicator "% recovery of phosphate from wastewater") allows distinguishing between current centralized solutions (where nutrients are normally not recovered) and decentralized options where nutrient recovery is often an explicitly stated objective (e.g. Larsen et al., 2009; 2012).

- Safe water supply and wastewater disposal → split into two fundamental objectives at the highest level: Good supply with water and safe wastewater disposal

  Reason: In the SWIP project, the wastewater infrastructure systems (C. Egger) are modeled separately from the drinking water infrastructure system (L. Scholten); the same applies to the MCDA for wastewater (J. Zheng) and water supply infrastructures (L. Scholten).

- Good supply with water / water for firefighting / good accessibility → deleted

  Reason: This attribute was characterized by the length of the hose to the building, which is obviously dependent on where fire hydrants are placed. We decided to base our dimensioning of the alternatives on the given current legal requirements for the case study utilities in the canton of Zürich, Switzerland (GVZ, 2011). The same applies to the current legal requirements for minimum water pressure (3.5 bar in the distribution system).

- Safe wastewater disposal / high reliability and / high service level of the drainage system → combined to one higher-level objective with two fundamental sub-objectives; the sub-objectives concerning no flooding in general / of less sensitive areas / and insensitive areas were deleted: Safe wastewater disposal / high reliability of the drainage system / few structural failures of drainage system and … / few overloads of drainage system

  Reason: The two sub-objectives concern the same objective, namely that one expects high reliability of the drainage system, i.e. that it does not block or collapse due to structural failures (leading to flooding), and that there are only few flooding under heavy storms. We use the following two attributes: For "few structural failures" we use the "weighted (by pipe diameter) number of pipe collapses and blockages / year /1'000 inhabitants"; weighting is done with the pipe diameter under the assumption that bigger pipes have a larger impact when they fail because more water is conveyed by them. Pipe failures are condition dependent and hence based on condition states predicted by a sewer deterioration model (Egger et al., 2013). For "few overloads of drainage system", we use the "weighted (by urban land use and number of inhabitants) number of incidents of insufficient drainage capacity per year (e.g. overflowing of manholes)" predicted by a hydraulic model. Here, we assumed that the damage is more severe if more people are affected, more dramatic in historic city centers, and the disturbance is higher if local trade or business is affected. Thus, we weighted this attribute by 1.5 if the area flooded is in a historic town center with mixed living and commercial zones.

- High social acceptance / high satisfaction of community and / high acceptance by end-users → deleted

  Reason: The hierarchical cluster distinguishing between the satisfaction (acceptance) of the community and the end-user is unnecessary and presumably only complicates elicitation, since it can, for example, also be important for the community to have low disturbance by road works. For similar reasons we removed all hierarchical clusters on the lower levels.

- High social acceptance / high satisfaction of community / high quality of management and operations / with three sub-objectives → sub-objectives deleted

  Reason: We decided to use the "% score of the EFQM Excellence model (European Foundation for Quality Management)" as attribute, since it is well-known and covers the relevant management aspects better than the sub-objectives that we invented. "The EFQM Excellence Model is the most popular quality tool in Europe, used by more than 30 000 organizations to improve performance", (EFQM, 2012). We asked an expert at Eawag (business economist) to classify our strategic decision alternatives accordingly.

- Low costs / easy fundraising → deleted

  Reason: Test-interviews for preference elicitation indicated that there are preferential overlaps between the two objectives "low annual costs" and "easy fundraising" because it proved difficult to get reliable estimates for real interest rates in the different alternatives. Additionally, we decided early in the project to not include financing strategies (e.g. are infrastructures fully financed via tax payers, are there subventions?). For these two reasons, "easy fundraising" was removed. As consequence, the real interest rate is also not considered in the calculation of the annual costs, but the discount rates still apply.

- Low costs / low cost fluctuations → reformulated as low cost increase

  Reason: We do not consider it as problematic if the costs decrease sometimes, while (large) increases are rather relevant.

5. Short discussion of objectives hierarchy and attributes

The construction of the objectives hierarchy was an extensive and careful process. First, we defined the system and e.g. decided that protecting floodplains is outside the infrastructures' system boundary. We included objectives that are often neglected in engineering practice and were judged less relevant by our stakeholders. These concerned social acceptance, future generations and the environment such as protecting groundwater (see above and Fig. 1, Tab. 1 in main text). We justify this to ensure that all pillars of sustainable development are included (Wuelser et al. 2012). Stakeholders tend to value current pressing issues higher than important solved ones from the past. For example, septic tanks were abolished in Switzerland due to groundwater pollution; groundwater quality is now high, and stakeholders judged groundwater protection as low priority. But for other cases and future generations, groundwater remains an important resource.

We need some objectives to distinguish between alternatives: "flexible system adaption" and "low unnecessary construction and road works" help to positively distinguish decentralized alternatives from the conventional central system, whereas "low time demand" and "low additional area demand for end users" are negative characteristics of these. Similarly, "high quality of management and operations" discriminates between organizational aspects. If this is not part of the decision, it can be excluded by giving it a scaling factor (weight) of zero. "Water for firefighting" might not be a requirement of the water supply system elsewhere. Other Switzerland-specific objectives might be "high autonomy concerning water resources" or "co-determination of citizens", since in many countries people cannot vote about (infrastructure) decisions.

We took great care to construct the attributes in such a way that they are applicable to other cases and that they comply with engineering requirements as well as decision theory. Some attributes may look similar, namely "few gastro-intestinal infections through direct contact with wastewater (due to failures of infrastructures)" and "few structural failures of drainage system" (Tab. 29; also see above and Fig. 1, Tab. 1 in main text). In both cases, the cause may be poor maintenance leading to collapses of pipes and back-pressure of wastewater into streets or cellars. However, the first objective refers directly to preserving human health; a fundamental goal of urban sanitation. The second aims at preventing the disturbance of daily business and traffic or the damage of property.

We regard this objectives hierarchy, as presented in the main text in Figure 1 and Table 1 to be exhaustive. It covers the main aspects important to water infrastructure planning. In application to other case studies, we recommend that those stakeholders carefully discuss which objectives are required for their specific decision situation and to delete those, which do not add additional insight. The attributes (Tab. 29) were constructed in a generalized way so that they are applicable to other cases. However, the respective ranges must be adapted to the boundaries in the respective application case, i.e. they should cover the worst- and best-possible decision alternative that is considered in that case.

Table 32: Description of the attributes that measure how well each objective is achieved. The short name refers to the objectives hierarchy given in Fig. 1 of the main text. We give the units, the ranges (worst- and best-possible state), the Status Quo, a more-detailed description of the attribute, and a narrative of the Status Quo in the case study region Mönchaltorfer Aa. DW = drinking water, WW = wastewater, WWTP = wastewater treatment plants, CSO = combined sewer overflows (discharge of mixed rain and wastewater without or with only basic treatment in the case of heavy rain events).

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
Intergenerational equity							
rehab	% Realization of the rehabilitation demand	[% realization]	0	DW: not completely realized WW: 80 - 100%	100 %	DW: 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 The realization of the rehabilitation demand for the period 2010-2050 is calculated as 1-[(no of failures per km) / (no of failures if nothing -except repair- is done)] * 100 % According to the recommendations of the Swiss Gas and Water Industry Association (SVGW), the failure rate should not exceed 0 1 failures per km WW: To keep the system as good as it is today, annual investments are needed These are approximately the reciprocal of the mean lifetime of pipes times the replacement value of the pipe network: Investment demand = (1[a]* replacement value) / (mean lifetime of pipes [a]) As an example, sewers have a lifespan of about 80 years To keep the system as good as it is today, about 1 25 % of the total system have to be rehabilitated every year: 1 / 80 years * 100 = 1 25 For each alternative, the effective investments in rehabilitation measures are summed up over the whole planning horizon and related to the total investment demand over the same period of time (also see Scheidegger et al , 2013)	DW: The rehabilitation demand is not completely realized (objective: <0 1 failures / (km*a), Status Quo ca 0 15 - 0 2 failures / (km*a) WW: Currently, 80 to 100 % of the total rehabilitation demand are being realized
adapt	Flexibility of technical extension or deconstruction of infrastructure	[% flexibility]	0	20 - 50 %	100 %	Expert assessment All alternatives were judged individually by 4 engineers according to how easy it is to technically extend or deconstruct the infrastructure The relevant aspects were: organizational structure, construction and operation of infrastructure, wastewater and drinking water system technology Each alternative was classified as: "very low (0 - 20 %)", "low (20 - 40 %)", "medium (40 - 60 %)", "high (60 - 80 %)", "very high (80 - 100 %) system flexibility" Using the mid-points of the intervals (10, 30, 50, 70, 90 %), the average and standard deviation were calculated Alternatives with >10 % deviation were discussed, and a final score assigned Larger interval ranges depict higher uncertainty or higher variance	Today's wastewater system is not very flexible (20 – 50 % flexibility) This is caused, amongst others, by the high path-dependency
Protection of water and other resources: Surface water							
chem	% Reference points in catchment that fulfill water quality target (nutrients, micropollutants, value > 0 6)	[% > 0 6]	0	50 %	100 %	Phosphorus in water bodies is an indicator of anthropogenic influences (via WWTP, CSOs, agriculture) and can lead to eutrophication In Switzerland, nitrogen is usually not a limiting factor for plant growth Nitrite is strongly toxic for fish Ammonium indicates pollution from wastewater or agriculture Dissolved organic carbon can be an indicator for anthropogenic pollution Total organic carbon includes particulate organic carbon, which reaches water after heavy rain from CSOs or organic fertilizers The biochemical oxygen demand is a measure for the oxygen used up by biological degradation processes; in severe cases, anaerobic conditions occur These can produce toxic substances as nitrite, methane, and hydrosulfides (source: FOEN, 2010) The Swiss Modular Concept for stream assessment is a new procedure to assess rivers and streams (Bundi et al , 2000, http://modul-stufen-konzept ch) To assess the chemical state, a set of nutrients are used (FOEN, 2010), and 3 indicators for pesticides that are relevant in the region Mönchaltorfer Aa (AWEL 2006) The nutrients are: total phosphor / (P_{tot}), total	Currently, the chemical state of the watercourse is "moderate" in the case study area Mönchaltorfer Aa 50 % of the reference points fulfill the water quality target level, based on a number of indicators for nutrients and pesticides For example, for three reference points, the concentrations of nitrate (NO_3) are higher than double of the target level, so that these reference points are judged as "very bad" concerning nitrate

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
						phosphor filtrated (P_{tot} filtr), orthophosphate (PO_4-P), total nitrogen (N_{tot}), nitrate (NO_3-N), nitrite (NO_2-N), ammonium (NH_4-N), total organic carbon (TOC), dissolved organic carbon (DOC), biochemical oxygen demand (BOD) The micropollutants are: photosynthesis inhibitors, chloroacetanilides, organophosphates For each substance, a target level is defined (concentration limits) The estimated level (from measurements or models) is compared with the target and classified (FOEN, 2010): "very good": estimated level of substance in watercourse is lower than half of the target level "good": estimated level is higher than half of target level and lower than target level "moderate": estimated level is higher than target level but lower than 1 5x target level "unsatisfactory": estimated level is higher than 1 5x target level but lower than 2x target level "bad": estimated level is as large as or even higher than 2x the target level To aggregate the results of each indicator at each reference point, we use an approach first described by Langhans and Reichert (2011) and Langhans et al (2013), which is further developed in the iWaQa project, based on multi-attribute value theory (Schuwirth et al 2012) The quality class obtained by each indicator is transferred to a neutral value between 0 and 1 with a value function The values are mathematically aggregated to give an overall assessment of the state of the watercourse We use a mix of additive and geometric aggregation, with equal weights for each indicator The reference points are existing measurement stations of (AWEL, 2006) with additional reference points added to the model used in iWaQa, basically upstream and downstream of urban areas To spatially aggregate the values at each reference point, we determine whether the estimated level is above the target If it is above, the water quality requirement is not reached (i e classes "bad", "unsatisfactory", "moderate" = value < 0 6) If the estimated level is below the target, the requirement is fulfilled (i e classes "good", "very good") Over the entire catchment, we give the % reference points that fulfill the quality requirements	
hydr	% Reference points in catchment that fulfill VSA guidelines for stormwater handling	[% yes]	0	44 - 74 %	100 %	The (VSA, 2002) guideline for a single discharge point evaluates the following relationship: V = Q_347/Q_E * fs * fg V = "Einleitverhältnis" = ratio between water amount coming from the river and water amount coming from the discharged rainwater [-] Q_347 = water flow in the river that is surpassed at 347 days a year (similar to the almost minimum water flow in the river) [m^3/d] Q_E = discharged rainwater flow after a rain event that occurs once a year [m^3/d] fS and fG = correction factors to account for the type of river and river bed Q_347 is derived from the model output of the water quality model of iWaQa and Q_E is determined from the total discharges from the combined and stormwater systems upstream of the individual reference points The result is evaluated in three classes: VG > 1: Discharge is allowed, only for very polluted water a treatment is required 0 1 < VG < 1: Discharge is allowed, but in water protection area, treatment is necessary VG < 0 1: Discharge is only allowed with prior retention The reference points are existing measurement stations of (AWEL, 2006) with additional reference points added to the model used in iWaQa (2012), basically upstream and downstream of urban areas To spatially aggregate the values at each reference point, we determine whether the VSA guidelines (2002) for stormwater handling are fulfilled or not Over the entire catchment, we give the % reference points that fulfill the guideline	Currently, 44 to 74 % of the discharge points fulfill the requirement of the VSA (2002) This means that in about half of the discharge points, the water that is led into the river can lead to turbulence distraction of the flora and fauna in the river
Protection of water and other resources: Groundwater							
gwhh	% Water abstraction / groundwater recharge	[%]	180	6	0	Will be presented in later paper by Lisa Scholten et al	

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
exfiltrsew	Water quality class (of nutrients; based on expert estimates)	5 classes	very bad	good	very good	One expert (Eawag scientist) classified the sewers according to the condition classes of (VSA, 2007), and another estimated the % wastewater that exfiltrates into the ground As indicators, we use the same nutrients as for the "good chemical state of watercourses" (see there), classified into one of 5 water quality classes (FOEN, 2010), but not the pesticides The concentration of each nutrient in wastewater is estimated based on average values from the literature (AWEL, 2006; GSchV, 2011; Gujer, 2002; Herlyn and Maurer, 2007) Then, the groundwater recharge rate is used to calculate the concentration of the nutrient in the groundwater For the condition classes according to (VSA, 2007), following % of wastewater exfiltrated was assessed with the experts: Class 0 (sewer is untight, has several cracks, is strongly incised, crushed, danger of collapse is given, floor is strongly corroded): 2 - 100 % of wastewater exfiltrates into the ground (average: 30 %) Class 1 (sewer is corroded or strongly eroded, has several cracks, has open pipe joints or some that broke off, loses water): 2 - 30 % of wastewater exfiltrates Class 2 (sewer shows damages, pipe joints are broken at the crown, some holes at the crown, has several cracks, that are sometimes strongly calcified, floor is slightly corroded and eroded): 0 - 15 % of wastewater exfiltrates Class 3 (sewer is in an insufficient condition The floor is slightly eroded, several small calcifications at the crown and the walls): 0 - 8 % of wastewater exfiltrates Class 4 (sewer is in a good condition): 0 - 4 % of wastewater exfiltrates Contrary to surface waters, there is no need to spatially aggregate the water quality classes at different reference points The groundwater body is regarded as an entity, and the calculations are based on the groundwater recharge rate of the entire system The influence of the soil (retention, degradation, hydraulic conductivity, height of the groundwater table) cannot be taken into account because there is not enough information This is why the uncertainty of this attribute is very high	Currently, the contamination through wastewater, for instance because of leaky sewers, is relatively low in the case study area Mönchaltorfer Aa (expert estimate: 8 - 10 %) The groundwater quality regarding nutrients is classified as "good"
exfiltrstruct	Water quality class (of biocides; based on expert estimates)	5 classes	very bad	very good	very good	The concentration of each biocide in the infiltration water is estimated based on average values from the literature (AWEL, 2006); Staufer and Ort 2012), and with an Eawag-expert Then, the groundwater recharge rate is used to calculate the concentration of the biocide in the groundwater Each biocide indicator is classified into a quality class, analogously to the "good chemical state of the watercourses" (see there) There are not a lot of nutrients present in infiltration water, so we did not consider these We only look at infiltration water from roofs (with Eawag expert) The influence of the soil (retention, degradation, hydraulic conductivity, height of the groundwater table) cannot be taken into account because there is not enough information available This is why the uncertainty of the attribute is very high Contrary to surface waters, there is no need to spatially aggregate the water quality classes at different reference points The groundwater body is regarded as an entity, and the calculations are based on the groundwater recharge rate of the entire system	Currently, there are no collection systems and infiltration structures to infiltrate water from roofs, parking lots, and streets in the case study area Mönchaltorfer Aa Thus, hardly any water from such areas that can contain biocides is being infiltrated Therefore, a "very good" water quality is assumed for infiltrated water
Protection of water and other resources: Efficient use of resources							
phosph	% Recovery of phosphate from wastewater	[% P recovery]	0	0	100 %	Phosphate recovery from urine is only done on laboratory and pilot scale at the moment With the current treatment it is possible to recover about 90 % of the phosphate (Etter and Kohn, 2007) Theoretically, it is possible to recover up to 100 %	Currently, no phosphate (as indicator for the recovery of nutrients) is recovered from wastewater
econs	Net energy consumption for water / wastewater treatment and transport	DW: [kWh / m^3] WW:	DW: 2 kWh / m^3 WW: 250 kWh / p /	DW: ca 0 5 kWh / m^3 (estimated) WW: 45 - 60	0	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 Status Quo was calculated/ estimated using data provided by the water supply / wastewater treatment plants in the case study area Mönchaltorfer Aa The worst case (maximum energy consumption) was calculated	Currently, the energy for treatment and transport of water is estimated to 0 5 kWh / m^3 (ca 46 kWh / person / year) in the case study region This equals about 0 25 % of the

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
		[kWh / p / yr]	yr	kWh / p / yr		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) With the gas produced during the digestion of the wastewater sludge, electricity can be produced using a gas-powered combined heat and power unit It is not only possible to produce electricity; heat can also be recovered from the wastewater stream with a heat exchanging device The heat energy is neglected because it only plays a minor role compared to electrical energy If inefficient use of electrical energy generates higher costs, these are considered separately in the objective "low costs"	energy requirement of a household, given current water usage For wastewater, the net energy for treatment and transport of wastewater in the central WWTPs of the case study area amounts to 45 to 60 kWh /person / year Compared to the total energy consumption of about 8'000 kWh /person / year, this equals about 0 6 % of the total energy requirement of a Swiss person (VSE 2012)
Good supply with water: Drinking water: Good quality							
aes_dw	Days per year with esthetic impairment such as taste, smell, etc	[d / yr]	365	0	0	Each alternative's esthetic water quality is assessed by an expert of the Cantonal Laboratory Zurich Details will be presented in later paper by Lisa Scholten et al	
faecal_dw	Days per year with hygienic concerns (hygiene indicators)	[d / yr]	365	0	0	Each alternative's esthetic water quality is assessed by an expert of the Cantonal Laboratory Zurich Details will be presented in later paper by Lisa Scholten et al	
cells_dw	Changes in total cell count as indicator of bacterial re-growth	[log]	2 (hundred-fold increase)	ca 0 68	0 (stable concen-tration)	Each alternative's hygienic water quality is assessed by an expert of the Department of Environmental Microbiology at Eawag and an expert of the Cantonal Laboratory Zurich Details will be presented in a later paper by Lisa Scholten et al ; also see (Lautenschlager et al , 2010)	Currently, there is approx a doubling of the cell counts after overnight stagnation
no3_dw	Anorganic substances (nitrate concentration)	[mg / L]	20	10	0	Will be presented in later paper by Lisa Scholten et al	
pest_dw	Pesticides (sum of pesticide concentration)	[µg / L]	0 15	0 036	0	Will be presented in later paper by Lisa Scholten et al	
bta_dw	Micropollutants (indicator: benzotriazole)	[ng / L]	150	105	0	Will be presented in later paper by Lisa Scholten et al	
Good supply with water: Drinking water: High quantity							
vol_dw	Days per year with water quantity limitations	[d / yr]	365	0	0	Will be presented in later paper by Lisa Scholten et al	
Good supply with water: Drinking water: High reliability							
ci_dw	Criticality index	-	0 25	estimated: 0 01 - 0 03	0	The criticality index is calculated as: criticality of affected pipe x probability of outage / total criticality of all pipes Will be presented in later paper by Lisa Scholten et al	
Good supply with water: Household water							
	Same objectives and attributes as "Drinking water"					Will be presented in later paper by Lisa Scholten et al	
Good supply with water: Water for firefighting							
vol_ffw	Available water for firefighting in new housing areas	[L / min]	500	ca 1'500	3'600	Will be presented in later paper by Lisa Scholten et al	
ci_ffw	Same as for "Drinking and Household water"					Will be presented in later paper by Lisa Scholten et al	

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
Safe wastewater disposal: Hygienic drainage and discharge							
illn	% Of total population getting infected once per year	[% / yr]	25 % / yr	0 001 - 2 3 % / yr	0 0002 % / yr	Wastewater contains human pathogens, but also from other sources (e g animal manure), if such wastewater drains into the sewer system (e g from farms) These pathogens can cause infections, which may lead to illness such as gastrointestinal disorders, especially in sensitive people (e g the elderly or children) Note that this risk is rather low We therefore use the % of the total population getting infected once a year as attribute If a person gets infected twice a year, he or she counts double in the calculation The attribute was calculated using the research of Ten Veldhuis et al (2010) A quantitative microbiological risk assessment is used to estimate the risk of illness due to exposure to micro-organisms after flood events and direct contact with wastewater For this, a dose response model for a certain infectious organism is required, which is combined with information about the exposure frequency The dose response models link the amount of a certain pathogen with the risk of infection at a single contact (P_single) There are many different models Ten Veldhuis et al (2010) use an exponential model for *Cryptosporidium* and *Giardia*, and a Beta Poisson dose-response model for *Campylobacter* The dose response models lead to very different results for different organisms in the same wastewater sample The risk of infection is therefore subject to a very high uncertainty Sampling: In (Ten Veldhuis et al , 2010) a series of samples was taken from combined sewers during dry weather flow *Cryptosporidium*, *Giardia*, *Campylobacter*, *E. coli*, and *Enterococci* concentrations were measured The *E. coli* and *Enterococci* concentrations found were compared with measurements of concentrations in flood water to roughly estimate the dilution during a flood event Based on this, the concentration of *Cryptosporidium*, *Giardia*, and *Campylobacter* in flood water could be calculated and then used in the microbial risk assessment (dilution factor: 10) With the exposure frequency (how many times does a person have contact with wastewater per year), an annual risk of infection can be calculated with: P_annual=1 - (1-P_single)^EF where P_annual is the annual risk of infection, P_single is the risk of infection per incident (result of the dose response model) and EF is the exposure frequency To define the amount of pathogens, a certain intake volume has to be defined According to the literature it was decided to use an intake volume of 10 to 30 ml per event The concentrations and the dose response models used were the same as in the work of Ten Veldhuis et al (2010) Exposure to wastewater may occur due to maintenance activities, failures, and flooding during extreme storms To estimate the predictions of this attribute for every alternative, the exposure frequency due to flooding will be defined by means of the hydraulic model), exposure due to failures with help of a failure model, and exposure due to maintenance due to literature values	The inhabitants of the region Mönchaltorfer Aa have direct contact with wastewater once every 10 years Between one person in 4 years (0 001 % of the population) and 547 people (2 3 % of the population) get infected with gastrointestinal pathogens every year (total population in region is 24'180 in 2011) Of those that get infected, ca 10 to 100 % get ill, depending on their body's defenses For the model, an average intake volume (20 ml) and an exposure frequency of 0 1 (contact with wastewater once every 10 years) was assumed
cso	Number of combined sewer overflows (CSOs) per year per receiving water	[no / yr / re-ceiving water]	60 CSOs / year / receiving water	10 CSOs / year / receiving water	plusminus 0 (0 001 = 1 in 100 years)	We know that currently up to 4 % of the population gets infected once per year with gastrointestinal pathogens after swimming or bathing in rivers or lakes This number is estimated with the average *E. coli* concentration at recreational sites in Switzerland and a model of the EPA (US Environment Protection Agency) for *E. coli* and gastrointestinal infection There is no information about CSOs underlying this approach, and we do not have any information for the case study region Mönchaltorfer Aa Therefore, we use the number of CSOs directly for this attribute Pathogens causing gastro-intestinal infections can also reach wastewater from agriculture, e g from animal manure It is usually not possible to distinguish whether the original source of infection is wastewater, or agriculture The worst case (maximum number of CSOs per year and receiving water) was defined by experience (Eawag scientist) The Status Quo was defined using the GEP ("genereller Entwässerungsplan"; urban drainage planning in Switzerland) for the town of Mönchaltorf This describes the number of CSOs into the river Mönchaltorfer Aa The best case is close to zero (1 overflow in 100 years) To make the predictions of this attribute for each alternative, we will use	In the year 2005, there were about 10 combined sewer overflows (CSOs) from the town Mönchaltorf into the river Mönchaltorfer Aa Hence, 10 overflows per year and per receiving water is considered to be the Status Quo

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
						the number of CSOs per year and receiving water, which are a direct output of the hydraulic models	
Safe wastewater disposal: High reliability of drainage system							
failure	Weighted (by pipe diameter) number of pipe collapses and blockages per year and 1'000 inhabitants	[no / yr / 1'000 people]	10 / yr / 1'000 p	0 0005 / yr / 1'000 p	0 0005 / yr / 1'000 p	Although this attribute seems similar to the ones above concerning "no gastrointestinal infections through direct / indirect contact with wastewater", it follows a different objective The previous ones refer directly to preserving human health This one refers ("only") to preventing nuisances, the disturbance of daily business, or the damage of property If a sewer is very large, it carries more rain and wastewater Consequently, if a larger sewer is damaged, there will also be a larger potential for wastewater being spilled into urban areas, and hence larger potential for damage than if the sewer is small We account for this by weighting the number of pipe collapses and blockages with the pipe diameter To estimate the range, the weighted pipe failure f was calculated as: f = l * r_f * g where l is the length of the sewer [km], r_f is the failure rate [/km/yr], and g is the weight: g= (D/D_average)^2, where D is the diameter of a certain pipe, and D_average is the average of all pipes of the sewer systems For the range, different failure rates were taken from the literature; minimum (for the best case): 0 0001/km/yr; maximum (for the worst case): 0 5/km/yr For two communities (Egg and Mönchaltdorf), an inventory of all pipes with their length, diameter, and location is given and used for the calculations To estimate the predictions of this attribute for each alternative, a model ("proportional hazard function") will be developed It links the condition class predicted by (Egger et al , 2013) to a failure rate	Today's drainage system is very reliable, we expect 0 0005 weighted pipe collapses and blockages per year and 1'000 people This equals one failure every 80 years in the case study region (24'180 inhabitants in 2011) In Mönchaltorf, for example, there are no reported failures In a bigger system, more failures can be expected As comparison, also in the city of Zürich there are hardly ever failures (confirmed by Zürich) The Zürich sewers are in very good condition and well maintained
service	Weighted (by city center and number of inhabitants) number of incidents of insufficient drainage capacity per year (e g overflowing of manholes)	[no / yr]	10 / yr	0 0002 - 0 13 / yr	0 (0 0002) / yr	This attribute may seem similar to the objective above "few structural failures of drainage system", because the final effects to the population might be similar, namely floodings of streets and houses with combined rain and wastewater However, we separate them, because they describe different types of troubles that are both important to urban drainage and wastewater engineers The causes for the attribute above are structural failures, and the prevention strategy is better maintenance and rehabilitation In the case of "sufficient drainage capacity", the causes are a too low hydraulic capacity of the drainage system, which can occur even if the system is very well maintained In this case, mitigation measures are the reduction of impervious areas (so that rain water drains directly into the ground), or can indirectly be addressed by planning the system differently (e g larger pipes and retention tanks, decentralized systems comprising larger retention and infiltration of stormwater) Of course, the nuisance or damage that such floodings cause is higher if more people are affected We weight the number of incidents by the number of inhabitants per hectare The damage is also more dramatic in historic city centers, and the disturbance is higher if also local trade or business are affected To account for this, we give a weight of 1 5 if the area flooded is in a historic town center with mixed living and commercial zones A 30 year historic rain series measured by a rain gauge located in the vicinity of the catchment area was used to evaluate the capability of the drainage system of properly draining stormwater For the worst case, it was assumed that no well-designed drainage system is present, so the water is mainly drained on surfaces and in trenches For the Status Quo, it was assumed that 20 % of the area is flooded every 10 years For the best case, it was assumed that the area is almost never flooded The damage d is then calculated as: d=(flooded area) / (total area) * flooding frequency * g Where g is the weight: g= (population density in flooded area) / (average population density) * 1 5 (for city center and mixed zones) The lowest weight is given to a zone with only single-family houses (a lot of area where water can drain off; e g big gardens), and the highest weight is given to residential and commercial zones with 4 story buildings	The drainage service is relatively high About 20 % of the area is flooded every 10 years due to insufficient capacity of the drainage system This leads to a weighted damage of 0 0002 to 0 13 per year, depending on the vulnerability of the flooded area (see "calculation attribute")

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
						To calculate the predictions of this attribute for each alternative, the frequency of overloading of each individual manhole will be calculated with a hydraulic model using historical rainfall series as model input To each manhole, an area is assigned which might be affected by flooding when overloading of the manhole occurs The area is characterized by the urban land use as indicator for its vulnerability to urban flooding The weight for the vulnerability can be by experts	
High social acceptance							
auton	% of the water coming from the region Mönchaltorfer Aa	[%]	0	55 %	100 %	The water supply from within and outside the case study region Mönchaltorfer Aa is calculated within the SWIP-project, based on the descriptions of each alternative and the water demand under the 4 future scenarios	On average, 55 % of the water comes from the case study region Mönchaltorfer Aa, and 45 % from lake Zürich
efqm	% Score of EFQM Excellence Model (European Foundation for Quality Management)	[%]	20 %	55 - 70 %	100 %	Each of the SWIP alternatives were assessed concerning their performance according to the EFQM Excellence Model (EFQM, 2012) by interviewing a business expert (Eawag scientist) The assessment is based on the organizational form and the geographic extent of our alternatives Through the 9 criteria of the EFQM Excellence Model, the firm can 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) The 9 criteria and their relative weightings are: 1 Leadership [10 %], 2 Strategy [10 %], 3 People [10 %], 4 Partnerships & Resources [10 %], 5 Processes, Products & Services [10 %], 6 Customer Results [15 %], 7 People Results [10 %], 8 Society Results [10 %], 9 Key Results [15 %]	The quality of management and operations under the current structures in the case study area Mönchaltorfer Aa can typically achieve 55 % to 70 % of the EFQM Excellence Model score, given favorable conditions
voice	Degree (percent) of co-determination	[%]	0	50 - 90 %	100 %	Each of the SWIP alternatives was assessed by two experts concerning the co-determination (Eawag scientists) They received documentation prior to the interview with a description of the relevant aspects for this attribute (organizational structure, geographic extent, financial strategy) As classification, the following semantic categories were used, and then translated into %: very low (0 - 20 % co-determination); low (20 - 40 %); medium (40 - 60 %); high (60 - 80 %); very high (80 - 100 %) In the case of differing estimates, the range was enlarged to cover both expert estimates This means that the lower % number was decreased, or the upper % increased As an example: if expert A gave an estimate from 40 - 60 % and expert B from 60 - 80 %, we used the total range from 40 - 80 %	Currently, the end users have medium to very high co-determination of about 50 - 90 % The system is a mix of responsibilities in the hands of households (household connections), cooperations, and the community The citizens are often directly involved in decisions by being able to participate in council meetings, or via public vote
time	Necessary time investment for operation and maintenance by end user	[h / person / yr]	DW / WW: 10 h / p / yr each	0	0	This attribute estimates the time each citizen has to invest per year to operate and maintain their decentralized water supply or wastewater disposal system This can involve e g the cleaning of filters, reading of meters, or the maintenance of tanks Also telephone calls to ask for help by a specialist, or complaints to a service hotline require time Estimates based on (realistic) times for maintenance of currently available decentralized (waste) water treatment units, and a number of telephone calls, based on expert estimates and product information	The current situation corresponds to the best-possible case Currently, there are practically no decentralized water supply or wastewater systems in the case study area Mönchaltorfer Aa that have to be maintained by the end users Hence, the time demand is 0 hours per person and year
area	Additional area demand on private property per end user	[m^2 / person]	DW / WW: 10 m^2 / p each	0	0	The range for this attribute was calculated using the area demand of decentral water or wastewater treatment units found in the literature (product information), and expert estimates Decentralized water supply systems cover the use of decentralized tanks with or without point-of-entry or point-of-use treatment In case of centralized supply, additional treatment can be installed in households One possibility for decentralized wastewater systems is a small treatment plant that works in the same way as a big central WWTP Another option is for example a septic tank, where the wastewater is stored before it is pumped out again and	Currently, there are practically no decentralized water and wastewater systems in the case study area Mönchaltorfer Aa that have to be installed on the private property of end users Apart from the installations for the pipes (including water meters and gate valves), the area demand thus corresponds to 0 m^2

Short	Attribute	Units	Worst	Status Quo	Best	Detailed description attribute and calculation	Status Quo
						transported away with a truck There are also low tech options such as constructed wetlands, which require the most area Hereby, the sewage water is lead into a planted field The plants take up the pollutants (e g nutrients) in the water and thereby clean it	
collab	Number of infrastructure sectors that collaborate in planning and construction	-	1	3	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 and the gas and water infrastructure could also soon need rehabilitation, these works could be carried out together Otherwise it could happen that right after the constructions works are closed by one sector, another sector starts its amelioration works, hereby reopening practically the same "hole"	Currently, in the case study area Mönchaltorfer Aa there is usually cooperation between the water supply and wastewater sector with the transportation department; i e 3 sectors collaborate In the community Gossau, for example, there are 2 joint meetings / year for planning and coordination In other communities there are joint meetings of road construction, water supplier, and wastewater utility as needed, i e if larger construction works are planned
Low costs							
costcap	Annual cost / person in % (DW) or in CHF (WW) of mean taxable income	DW: [% / p / yr] WW: [CHF / p / yr]	DW: 5 % / p / yr WW: 863 CHF / p / yr	DW: 0 4 % / p / yr WW: 289 CHF / p / yr	DW: 0 01 % / p / yr WW: 76 CHF / p / yr	For wastewater, the calculations for the range are based on numbers in a report of (VSA, 2011); the Association of Swiss wastewater and water protection experts Hereby, all Swiss communities were asked to provide their cost data In the (VSA, 2011) report, the total annual costs consist of running and capital costs The running costs consist of the labor and material costs The capital costs consist of the imputed depreciation costs and the interest costs The transport costs for sludge transport is included for decentralized treatment options The money needed for the water supply and wastewater infrastructure can be collected in numerous forms through taxes, tariffs, and direct payments, which we do not consider For the water supply sector, we decided to elicit cost-preferences as percentage relative to the mean taxable income (65'000 CHF / p / yr for federal taxes, averaged over the four communities in the area of the case study Mönchaltorfer Aa) For the wastewater sector, we decided to elicit the preferences by using the annual cost in CHF per person to measure this attribute The detailed cost calculations for each alternative will be carried out by an engineering company	Currently, the total costs for water supply in the region Mönchaltorfer Aa amount to ca 0 4 % (273 CHF / p / yr) of the average taxable income (ca 65'000 CHF / p / yr) The total costs for the entire wastewater disposal system amount to 289 CHF per person and year, based on the average total annual costs of wastewater treatment plant and the sewer system for the year 2011
costchange	Mean annual linear increase of costs in% (DW) / in CHF (WW) per person and year until 2050	DW: [% / p / yr] WW: [CHF / p / yr]	DW: 20 % / p / yr WW: 43 CHF / p / yr	DW: 8 % / p / yr WW: 1 4 CHF / p / yr	0	To estimate this attribute, the total annual costs will be calculated for every year (see attribute "low annual costs") The increase of costs from 2010 to 2050 will be divided through 40 and averaged for the cost increase per year	In the case study area Mönchaltorfer Aa, the total costs for water supply from 2006 to 2010 increased on average by 8 % (linear increase) For wastewater disposal, the costs have increased by 1 4 CHF per person per year in the last five years (20'864 CHF higher costs / year at an average running cost of 776'975 CHF / year) For wastewater, we use accounting information about the running costs of the WWTP in the case study area Mönchaltorfer Aa from 2006 to 2010

A2 Future scenarios

1. Methods scenario workshop

A preliminary objectives hierarchy was created on the desktop by the project team and discussed with the stakeholders in the 27 face-to-face interviews (Figure 1; also see Lienert et al. 2014). Details concerning this interview series and the stakeholder selection are also given in the stakeholder and social network analysis (Lienert et al. 2013).

Three future socio-demographic scenarios for the case study region for the year 2050 were created in a stakeholder workshop in April 2011. 15 of 22 invited participants from the case study region participated. After a general introduction to the project and the ideas behind scenario planning, we presented three scenarios that differed in six main characteristics: global situation, environment, spatial development, population, working, and transportation. Furthermore, we presented eight factors that characterize the water supply and wastewater system: quantity of water used and wastewater generated by the population, quantity for industry, societal requirements concerning water quality, legal requirements concerning drinking water and wastewater treatment, spatial development of the communities, financial situation of the communities, financial situation of population and industry, and subventions and tax incentives. The factors were discussed in groups of two and then in the plenum to eliminate factors that are not relevant for the region or to include other very important factors.

We then assigned the participants to three groups with mixed stakeholder types and assigned a scenario to each group. Each group discussed what the general development in 2050 could mean for their communities, and they were asked to conjure a vivid, detailed, and coherent picture. In the next step, they were asked to describe in detail how the water supply and wastewater system might look like in the respective future world; they were asked to be as specific as possible and to use numbers (e.g. for population growth or water consumption). The scenario specification was based on the factors that had been previously discussed and modified in the plenum (Tab. 31). They chose a title for their scenario, noted the core characteristics on a flip-chart, and made a sketch to visualize the main ideas. The three scenarios were presented in the plenum.

2. Results scenario workshop

The eight factors that characterize the water supply and wastewater system were discussed in the plenum. One factor was eliminated by merging (financial situation of population and industry merged with financial situation of community), and three were added: coordination among the communities, environmental impacts, and availability of energy (Tab. 31). The factors "availability of resources and materials" and "available technologies" were discussed in the plenum but not included in the list of mandatory factors. However, the groups could include them if they wished.

Three future socio-economic scenarios were created in the groups (details in Tab. 32). Note, that we later modified certain characteristics defined in the scenario workshops; namely the spatial planning in the "Boomtown Zürich Oberland" scenario" and the water demand per person and day (also see Lienert et al. 2014).

Table 33: Factors to construct scenarios. Description of the factors that describe the water sup-ply and wastewater system, which were given to the workshop participants, discussed and adapted in the plenum, and finally used to specify the future scenarios created in three stakeholder groups. WWTP = wastewater treatment plant.

	Factor	Description
A	**Quantity** of water used and wastewater generated by the population	Describes two developments: (i) the demographic development (i.e. population growth) and (ii) the specific water demand of households. Will future lifestyle change the required water quantity? We assume that the wastewater quantity is similar to the supplied water quantity.
B	**Quantity** of water used and wastewater generated by the industry	Describes the requirements of industries that are relevant for water management. The water demand and wastewater production (especially the load of contaminants) should be described separately.
C	Societal requirements concerning **water quality**	What services do the people and consumers ask from the urban water management system? For example, are they very environmentally-friendly and health-conscious and would they also be willing to pay more for water and wastewater treatment than required by law? Would they also pay for the elimination of micropollutants in drinking water or for the hygienization of the wastewater overflows from WWTP?
D	**Legal** requirements concerning drinking water and wastewater treatment	Describes the legal requirements and norms for water supply and wastewater treatment. As an example, is it required by law to monitor a number of micropollutants in drinking water and to remove these? Are there more stringent requirements for wastewater treatment such as the hygienization of wastewater overflows from the WWTP? What are the requirements for firefighting?
E	**Spatial development** of the communities	Describes the type of settlements and the building activities in the communities. Will there be densification or urban sprawl? Will there be mainly apartment houses or single-family houses? Where will there be buildings and where not?
F	**Financial situation** of the communities (and population, industry)	Describes the financial degrees of freedom and the possibilities of the communities, population, and industries in the region. Are these heavily indebted? Is there sufficient public (tax) money available?
G	*Financial situation of population and industry*	*Merged with F after discussion in the plenum.*
H	**Subventions** and tax incentives	How is the urban water management system financed (e.g. with public tax)? Are there tax incentives (e.g. wastewater bills, taxes to deal with water shortages or to remove micropollutants)? Are there subventions (e.g. to hygienize the outflow from WWTP for re-use in agriculture)?
I	**Coordination** among the communities	Describes how the communities are organized. Is there a separate political and management system for each community? Do the communities collaborate (and if yes, how)? Are there mergers of communities into one larger entity?
J	**Environmental** impacts	For example, consequences due to depleting water resources. Consequences of activities in the region and the water infrastructures on the quality of water bodies.
K	Availability of **energy**	For example, what are the consequences of energy shortages for the water sector? Is energy generated and/ or stored by the water supply and wastewater system?

Table 34: Scenario description. Description of the three socio-economic scenarios for the case study region near Zürich that were created in a stakeholder workshop: (A): Boom, (B): Doom, (C): Quality of life. WWTP = wastewater treatment plant, CSO = combined sewer overflow (mixed rain and wastewater is discharged directly to rivers and lakes without treatment or only very basic treatment in the case of heavy rain events).

	Scenario	General situation	Spatial development	Transportation	Financial situation	Collaboration	Water supply	Wastewater system	Energy and environment
A	Boomtown Zürich Oberland (Silicon Valley Aabach)	In 2050, Europe belongs to the most prosperous regions worldwide Region Mönchaltorfer Aa is booming Massive population growth from today 25'000 inhabitants to ca 200'000 High-tech industries with high productivity; large trust in technologies	Region is very densely populated High land prices; very dense urban development (25-story-buildings) Few villas for the rich Few agricultural areas and nature protection zones Recreational areas (river Aabach, lake Greifensee)	Strong increase in mobility; commuters from E-Switzerland New transport axes (highway, access roads, magnetic levitation train)	Communities prosper, rising tax revenues Loans for infrastructure investments needed, but also higher income (more connection fees) No subventions, financing only via fees Tax incentives foster use of water of different qualities	Communities are forced to collaborate due to high dynamics in region	Overall increase of water demand (population growth), but considerably lower per person water demand due to clean-tech Some areas distribute water of different quality (drinking water, household water, firefighting) No shortages due to access to lake Zürich High water quality standards promote closed-loop technologies and on-site treatment Health-consciousness of people leads to high requirements for drinking water quality (at least as good as today)	Central WWTP in industrial zone, mainly for household water (no heavy industries) Much stricter requirements for wastewater treatment to compensate population growth Remaining nature protection zones (Aabach, lake Greifensee) similarly clean as today (no smell or eutrophication) Additionally, removal of micropollutants is required from society and by law Climate change leads to heavy rain events and various measures for discharge management and flood control	Environmental protection and quality of life very important High costs for fossil fuels: resource stewardship; use of renewables Per person energy consumption much lower (clean tech)
B	Doom	Increasing gap between Europe and prospering Asia Switzerland is increasingly unattractive in the global world This causes strong financial pressure on public provisions, especially of infrastructures in water sector that have high investment costs Decline of industries Deregulation	Spatial development of communities stagnates Relatively strong urban sprawl Slight population decline		Despite high investment needs and rising costs it is politically not possible to raise water fees No subventions or state finances	Increased collaboration between communities to make use of synergies and expertise	Water demand decreases to 80 l / person / day (ca 2x less than today)[a] Communities reduce capacities and investments Very bad state of pipes Strong dependence on local water sources; highly variable quality (on average only household water) Hence, population has own water sources; e g bottled water, rain water collection (garden, membrane filter for kitchen) Control and monitoring by state hardly existent and ineffective Drinking water quality standards as 2011, but not relevant (bottled water etc) Minimal requirements for fire water	Wastewater quantity is lower by ca 25 % than in 2011 Negligible inputs from industries Separate sewers for wastewater only are abandoned; only mixed sewers (rain and wastewater together) Climate change effects are strongly perceptible in urban drainage with increasing floodings after heavy rain events and increasing CSOs WWTP are in a very bad state They are held together with "spit and tape", with frequent failures Only mechanical parts are functioning reliably Lower wastewater quality standards	Environmental effects (deficient wastewater treatment; climate change (CSOs)) Decreasing concern about micropollutants Energy is expensive (saved wherever possible)
C	Quality of life	Europe belongs to the most prosperous regions In Europe, Switzerland is important Moderate population growth (<5 % / year; 20 % until 2050) High environmental and health awareness High productivity in agriculture; high ecological standards	Additionally required residential areas mainly created by more dense urban development, rather than providing more land for buildings Only 5 % additional building areas (= ca today's reserves of building zones)	Public transport is promoted and efficient Commuting is reduced by actively promoting e-technologies (ca 30 % home office)	Financial situation of communities and population is good Sufficient finances for good maintenance and operation of the water infrastructures available	Grüningen and Gossau are merged Mergers with other communities are discussed, following general trend in ct Zürich: 50 communities in 2050 (2010: 171)	Higher drinking water quality (sensitive analytics; better information about chronic effects) Water demand of households lower than today (140–150 L / person / d)[a]; of industry as 2011; higher in agriculture Water supply by public network, rain water retention basins, and advanced treatment ponds Technical requirements for network are lower Cost savings due to smaller pipe diameters and new laying techniques Flexible fire water provision, coupled with rain retention measures	Very high quality requirements for wastewater treatment, and protection of the environment and water resources Discharge from WWTP reaches nearly drinking water quality standards Depleting resources, high energy prices, and climate change effects have led to constant optimization and new developments E g nutrients are recycled from wastewater and used as fertilizer in agriculture	Very high environmental standard; resources recycling Energy production from biomass; energetic optimization of wastewater system

[a] We could not directly use the water demands specified in the workshop; Lienert et al. (2014).

A3 Step (3) Develop alternatives

1. Methods workshop to create alternatives

In the 2nd stakeholder workshop in May 2011, the twenty participants created strategic alternatives with help of a strategy generation table. We prepared the 17 factors and their specifications beforehand. The 17 factors concerned the organizational structure (four factors; e.g. cooperation between sectors), geographic extent (two factors; e.g. cooperation between communities), financial strategy (two factors; e.g. rehabilitation strategy), construction and operation of water infrastructure (four factors; e.g. operation & maintenance), wastewater system technology (two factors, e.g. storm water handling), and drinking water system technology (three factors, e.g. central or decentralized water treatment). The strategy generation table is given in Table 33.

The participating stakeholders were split into four groups according to their professional background. We mixed groups to ensure the representation of different perspectives (local, cantonal, and federal stakeholders, and actors from different sectors, i.e. water supply, wastewater, administration). Each group was assigned to one of the four change scenarios specified in the first stakeholder workshop (Boom, Doom, Quality of Life, and Status Quo; Tab. 32). We asked the participants to create at least two different alternatives per group. First ideas of possible alternatives were collected by each group during a 15 minute brainstorming under the premises of the assigned change scenario. Each group then selected some of the generated alternatives (the favorite one, the most probable one, etc.), which was further systematically characterized by choosing (or generating new) specifications of each factor from the strategy generation table (Tab. 33). Some of the factor specification required a more-detailed definition. As an example, for the funding strategy (factor G, Tab. 33), specifications c) and d) required numbers concerning the % self-financed in the constant budget, or the % increase per year in the progressive budget. The most important characteristics were presented in the plenum. Altogether ten decision alternatives were defined. The project team used these backbones as input to develop more-detailed alternatives to be used in the later MCDA.

Table 35: Strategy generation table. Overview of 17 factors (A – Q) and the respective factor specifications (a – h) in six main categories: Organizational structure, geographic extent, financial strategy, construction and operation of the infrastructure system, and system technology of the wastewater and drinking water system. DW = drinking water, WW = wastewater, WWTP = wastewater treatment plant.

Organizational structure				**Geographic extent**		**Financial strategy**		**Construction and operation of water infrastructure**				**Wastewater technology**		**Drinking water system technology**		
A Form of organization	**B [a] Cooperation sectors: DW, WW, others**	**C [a, b] Responsibilities WW sector**	**D [a, c] Responsibilities DW sector**	**E Cooperation communities**	**F Cooperation w. other communities**	**G Funding**	**H Rehabilitation strategy (DW & WW)**	**I Rehabilitation measures**	**J Pipe / sewer laying technique**	**K Operation & maintenance**	**L Inspection & surveillance**	**M Drainage system**	**N Storm water handling**	**O Purpose of use**	**P Distribution system**	**Q Water treatment**
a) Community	a) DW / WW / others	a) Private / sewer / WWTP	a) Intake / treatm / distr / private	a) All individually	a) None	a) Constant budget, 100 % self-financed	a) Rehabilita-tion of x % of network	a) Replace	a) In trench	a) Extensive	a) A lot (to be defined)	a) Combined sewer (1 sewer)	a) Discharge	a) Water for food (drinking & cooking)	a) Centralized	a) Centralized (to be defined)
b) Cooper-atives	b) [DW + WW] / others	b) Private / [sewer + WWTP]	b) [Intake + treatm] / distr / private	b) 2 together, the others individually	b) Wetzikon	b) Constant budget, 0 % self-financed	b) Condition-dependent measures	b) Repair	b) Trenchless	b) Moderate	b) Average (to be defined)	b) Separate (2 or more sewers)	b) Retention	b) Water for hygiene (e g shower)	b) Decentral-ized tanks (e g roof)	b) Decentral-ized (to be defined)
c) Operator model: franchising	c) [DW + others] / WW	c) [Private + sewer] / WWTP	c) Intake / [treatm + distr + private]	c) 3 together, 1 of others individually	c) Uster	c) Constant budget, x % self-financed	c) Rehabili-tation basis = prioritization	c) Renovate		c) Minimal	c) Little (to be defined)	c) Decentral-ized	c) Infiltration	c) Water for cleaning & garden	c) Super-market (bottles)	c) Combin-ations
d) Operator model: contracting	d) [WW + others] / DW	d) [Private + sewer + WWTP]	d) [Intake + treatm + distr] / private	d) All 4 together	d) Maur	d) Progress (x % annual ncrease)	d) Measures only upon urgent need	d) Do nothing		d) Do nothing	d) None at all	e) Semi-(de-)central-ized	d) Combin-ations	d) Water for fire fighting	d) Delivery service (tanks or bottles)	d) None at all
e) IKA = inter-communal agency	e) [DW + WW + others]		e) [Intake + treatm] / [distr + private]	e) Parts of communities with a) – e)	e) Whole Greifensee area including Pfäffikersee		e) Do nothing					f) Combin-ations		e) Water for emergency supply	e) Decentral-ized ponds	
f) Corporation			e) [Intake + treatm + distr + private]		f) Whole Gr see excl Pfäffikersee										f) Household delivery from community	

Organizational structure				Geographic extent		Financial strategy		Construction and operation of water infrastructure				Wastewater technology		Drinking water system technology		
A **Form of organization**	**B** [a] **Cooperation sectors: DW, WW, others**	**C** [a, b] **Responsibilities WW sector**	**D** [a, c] **Responsibilities DW sector**	**E** **Cooperation communities**	**F** **Cooperation w. other communities**	**G** **Funding**	**H** **Rehabilitation strategy (DW & WW)**	**I** **Rehabilitation measures**	**J** **Pipe / sewer laying technique**	**K** **Operation & maintenance**	**L** **Inspection & surveillance**	**M** **Drainage system**	**N** **Storm water handling**	**O** **Purpose of use**	**P** **Distribution system**	**Q** **Water treatment**
g) House-holds					g) City of Zürich										g) Combin-ations	
					h) Region Zürich Oberland											

[a] Interpretation for B, C, and D: as an example, [DW + WW] / others means that the drinking water and wastewater infrastructures are managed together by one entity, while other infrastructures (e g electricity, gas supply, telecommunication) are separately operated by another entity

[b] Here, “private” mean the house drainage sewer pipes on private ground

[c] Here, “private” means household connections for water supply on private ground; “intake” means retrieving water from a source; “treatm” refers to the drinking water treatment; “distr” refers to the distribution and storage of drinking water

2. Results strategic decision alternatives

In the second stakeholder workshop in May 2011, the twenty participants created strategic alternatives with help of a strategy generation table. We prepared the 17 factors and their specifications beforehand. The 17 factors concerned the organizational structure (four factors; e.g. cooperation between sectors), geographic extent (two factors; e.g. cooperation be-tween communities), financial strategy (two factors; e.g. rehabilitation strategy), construction and operation of water infrastructure (four factors; e.g. operation & maintenance), wastewater system technology (two factors, e.g. storm water handling) and drinking water system technology (three factors, e.g. central or decentralized water treatment). The strategy generation table is given in Table 32.

The specifications of each factor for each of the 10 alternatives that were created in the stakeholder workshop are summarized in Table 33. The alternatives were then processed by the research project team to ensure internal consistency. Moreover, to better describe alternatives, we created following additional factors: water source, water treatment, operations, technical planning, administration and support, leadership, strategy, and partnership and resources. Some factors were necessary to predict the objective "high quality of management and operations", for which we used the attribute "% score of the EFQM model" ("The EFQM Excellence Model is the most popular quality tool in Europe, used by more than 30'000 organizations to improve performance"; EFQM 2012). The more-detailed description developed by the project team, is given in Table 34. In the following we give the narratives for each strategic decision alternative, based on the workshop results. Note that the alternatives were created having a certain future scenario for the year 2050 in mind but that they will be evaluated in the MCDA for their performance under all four future scenarios. Following a recommendation of Gregory et al. (2012a) we re-named the alternatives so that their names are better understandable also to those that did not participate in the workshop.

Alternatives for scenario A, "Boom"

A1a) Centralized, privatization, high environmental protection

All network infrastructures are combined together (water, wastewater, gas, roads, telecommunication, and electricity) and managed by one private single entity that charges fixed fees for its services. Whereas sophisticated contracting is necessary, conflicts of interest arise between the municipalities, the wider public, and the contractor. Maintenance is mostly asset-related. New buildings are mainly equipped by green rooftops for stormwater retention.

A1b) Centralized, IKA

Differs from variant 1a) only in the fact that an intercommunal agency (IKA, "Interkommunale Anstalt") manages the infrastructure, not a contractor.

A2) Centralized, IKA, rain stored

Although combined, the wastewater, drinking water, and gas infrastructure services remain in the public domain, but private sector principles rule their management. Their maintenance is asset-related and pipe or sewer laying is done in the most economic manner. No dedicated retention of stormwater is foreseen.

A3) Fully decentralized

The water infrastructures are as decentralized and as much reliant on the consumers as possible. The responsibility for the water and wastewater service is privately owned so that the centralized infrastructure is minimal. Storm water is collected, reused, and infiltrated where possible.

Alternatives for scenario B, "Doom"

A4) Decaying infrastructure; decentralized outskirts

Water infrastructures are still centralized, but local sector combinations exist. Outside current residential areas, the communities have transferred the responsibility for sewerage and storm water management to the private consumer. The currently existing wastewater system is still publicly operated, while newly developed areas are not served by a well-designed buried sewer system. Instead, stormwater from these areas is simply drained on the surface of roads and via trenches and sanitary wastewater is treated in septic tanks. The existing central WWTPs decay and provide only mechanical treatment. The quality of the piped water supply is not apt for drinking (no treatment). Consumers buy their water for food in the supermarkets or have their own household treatment. No real budget is available. Whenever funding is available, it is allocated in the most sensible way. Consequently, maintenance and inspections are only performed based on importance classification of the pipes and sewers. Rehabilitation only takes place if at least 100 consumers are affected, otherwise only repairs will be done.

A5) Decaying infrastructure everywhere

Most infrastructure services as well as their funding are in the responsibility of the customers. In general, no public funding is available anymore for the maintenance of the distribution, collection, and treatment systems. Therefore, wastewater is technically managed as in A4. However, sludge from septic tanks is collected privately.. There is no centralized water supply, and no more pipes are being built. Consumers are accountable for their own water supply and operate tanks which are intermittently recharged. Water is delivered to the households by a private delivery service and treated in house (e.g. with activated carbon). The municipalities – or parts of them – are partially combined. Operational and maintenance efforts are considerable where affordable, but then again no inspection and surveillance are done.

Alternatives for scenario C, "Quality of life"

A6) Maximal collaboration, centralized

One of the main ideas behind this alternative is to increase the decentralized use of rain water in the households and provide considerable retention volume under intensive rainfalls. Despite this, centralized drinking water supply and drainage remain. Only about 10 % of the drinking water (mainly surface water from the lake) is treated. The service provider of the four case study communities and Oetwil am See is a cooperative that combines the water and wastewater services with telecommunication, electricity, gas, and road services. A 100 % self-financed constant budget is available for the realization of rehabilitation measures according to the condition of the infrastructure. Efforts for operation and maintenance, as well as inspection intervals are neither low nor high.

A7) Mixed responsibility, fully decentralized with onsite treatment

Public water supply and wastewater services are combined within one cooperative for all 4 case study communities. However, treatment facilities on private grounds (households, industry) are within the responsibility of the owner. A central wastewater treatment plant and centralized storm water sewers are operated by the cooperative, but no sanitary wastewater sewers will be constructed in new development areas. Stormwater is retained as extensively as in A6. The water infrastructure is mostly decentralized, with on-site drinking water treatment and wastewater treatment with urine source separation for nutrient recovery. This fertilizer finds its use in local agriculture. The water demand within households is strongly reduced thanks to modern vacuum toilets. The concentration of wastewater is thus high. Water within the households is reused as far as possible (especially rain water) and is only delivered (with tank trucks) by the municipality upon special demand or in longer dry periods. The firefighting policy is based on fire engines that withdraw firefighting water from central water storage ponds. All residues (e.g. sludge) from on-site treatment installations are transported by truck to central treatment and disposal facilities. Rehabilitation of the infrastructure is 100 % self-financed and prioritization is according to condition. Operation and inspection efforts are medium, as in alternative 7. The infrastructure organization, structure, and management in the surrounding urban areas are comparable.

Alternatives for scenario "Status Quo"

A8a-e) Status Quo with storm water retention (drinking water only 8a-b)

While the communities remain responsible for a single, integrated wastewater and drinking water sector, some services are contracted out to private enterprises. The water infrastructures of Egg, Gossau, Grüningen, and Mönchaltorf are jointly operated and maintained. Funding is flexible owing to a mix of 50 % leverage and 50 % self-finance. The quality of construction and maintenance is high and regular inspections lead to a good comprehension of the underground infrastructure. The standards and legal requirements are respected, and the STORM guideline (VSA 2007b) is widely implemented. To prioritize the development of the wastewater system, the Swiss water protection law (article 7, Abs. 2, GSchG) is interpreted as follows: 1st infiltration of storm water, 2nd separate sewer system (storm water is discharged to surface water bodies, if possible following retention or treatment), and 3rd combined sewer system. While the capacity of the sewer network remains the same as today (2011), optimization in wastewater treatment leads to higher quality of the treated wastewater. Water for food, hygiene, cleaning, and firefighting is distributed through a pipe network from a central treatment facility, as today. Several variants of this alternative are elaborated comprising decentralized as well as centralized treatment options at different locations and scales of the wastewater system.

A9) Centralized, privatization, minimal maintenance

This alternative reflects how the stakeholders believe that an unfavorable development under current conditions could look like. It differs from alternative 8 (Status Quo) mainly with regard to organization, finance, and maintenance while the legal framework and technical wastewater and drinking systems are roughly the same. Due to privatization, consumers can choose their water service provider (e.g. water from a supermarket provider; in general all providers seek for revenue-maximization). Funding is 100 % leverage-based and despite rising fees, no financial sustainability is obtained. This is partly due to the fact that rehabilitation measures are only undertaken when urgently needed. The efforts for operation, maintenance, and inspection of the water infrastructure network are also minimal. The horizontal (sectoral) as well as vertical (no merging of communities) fragmentation of 2011 remain (see Lienert et al. 2013).

Table 36: Alternatives composition matrix. Factor specifications of the ten (nine and two variants of alternative 1) strategic decision alternatives that were created in the stakeholder work-shop. Columns represent factors, rows the chosen factor specifications. Each number (1–9) represents one alternative. Shaded fields indicate factors that were specified beforehand (as in Tab. 7); fields with numbers but without shading indicate that a new specification was created by the workshop participants for the respective alternative. Empty shaded fields were not chosen. Fields with blue shading indicate factor specifications that were removed by the participants. Reading example: Alternative A4 ("Decaying infrastructure; decentralized outskirts") consists of the factor specifications: A a) [or b) or g)], B a), C b), D g), E e), F i), G e), H d), I b), J a), K c), L c), M e), N c), O b [or c)], P g), Q d). DW = drinking water, WW = wastewater.

	A Form organ-ization	B Coop-erat. sector	C Res-pons. WW	D Res-pons.DW	E Coop. comm-unities	F Coop-erat. others	G Fund-ing	H Rehab. strat-egy	I Rehab. meas-ures	J Pipe / sewer laying	K Oper-ation,maint.	L Ins-pect-ion	M Drain-age syst.	N Storm water	O Purp-ose use	P Distri-bution syst.	Q Water treat-ment
a)	4, 8	3, 4, 7, 9	9	9	(3), 9	2, 3 8, 9	2, 6 7		5, 6	4, 5, 6, 7, 9	1, 5				1,2,5,6, 7, 8, 9	1, 2, 6, 8, 9	1, 2, 3, 6, 8, 9
b)	4, 6, 7	5, 8	4, 5, 6				9	2, 6	1, 2, 3, 4, 6, 9	8	2, 6, 7, 8	1, 2, 6, 7, 8	1, 2, 6, 7, 8, 9	6, 7	3, 4, 5, 6	3, 5, 7	5, 7
c)			7				8	1, 7, 8	7, 8	1, 2	3, 4, 9	3, 4, 9	3, 7	4, 5, 8, 9	4, 5	3	
d)	1a, 9		1, 2, 8	6	1, 2, 6, 7, 8			3, 4, 5, 9				5	5	1, 2, 3, 5		3, 5	4, 6
e)	1b, 2	1, 2, 6	3		4, 5		1, 3, 4, 5						4			3	
f)				1, 2, 7, 8													
g)	3, 4, 5, 7, 9			4, 5												4	
h)						1											
i)						4, 5, 6, 7											

Table 37: Definition of strategic decision alternatives. Overview of nine strategic decision alternatives (and some variants) for water supply and wastewater infrastructures in the study region Mönchaltorfer Aa. The alternatives were initially developed in a stakeholder workshop and thereafter processed by the research project team to ensure internal consistency. For simplicity, we grouped the 17 factors (Tables 33,34) and provide a general description together for: organizational structure, sector cooperation, and management (factors A–G), rehabilitation strategies, operation and maintenance (factors H–L), and wastewater and water supply system technology (factors M–Q). WW = wastewater, WWTP = wastewater treatment plant.

No.	Alternative name	Organizational structure, sector cooperation, management	Rehabilitation strategies, operation, and maintenance	Wastewater and water supply system technology
A1a	Centralized, privatization, high environmental protection	One private organization manages all sectors [(a)] and all communities [(b)] together (also with entire region Zürich Oberland) Equal partnership with contractor who charges fixed fees Performance-based leadership that achieves promised service levels at minimal costs	Rehabilitation is done according to prioritization [(c)] Decision about measures (replace, repair) are related to asset The extensive operation and maintenance is comfortably performed through underground service galleries, but inspection is only average	The water supply and wastewater system are fully centralized Large amounts of water are supplied in drinking water quality, and can also be used for firefighting There is a 4th treatment step at the WWTP to remove micropollutants New developments outside existing building zones are drained by separate systems New houses are equipped with green rooftops for retention of stormwater
A1b	Centralized, IKA	Differs from A1a only in the fact that an intercommunal agency (IKA) manages the infrastructure, not a contractor Technocrat leadership (very experienced and qualified, but rather rigid) with focus on maximizing performance	As A1a	As A1a
A2	Centralized, IKA, rain stored	As A1b, but constant budget, 100 % self-financed	Rehabilitation is done according to condition [(d)] The decision about measures (replace, repair) is related to asset, and the most economical pipe laying technique is used Their operation, maintenance, and inspection are only moderate	The water supply and wastewater system are fully centralized, as A1a However, water for firefighting is only partially supplied through the network, and is gained as far as possible from rain water, which is retained in underground firefighting tanks No dedicated retention of stormwater foreseen
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 organizations that have a long-term relationship with their customers	Only repairs, but no rehabilitation is undertaken, and only upon urgent need for action Operation and maintenance are moderate, while there is little inspection	The water infrastructures are as decentralized as possible, only minimal centralized infrastructure Storm water is collected in households, decentrally treated and reused for household water and firefighting Drinking water is bought in the supermarket Gray water is treated locally and fed into water supply tank, rest is treated centrally Excess storm water is wherever possible infiltrated
A4	Decaying infrastructure; decentralized outskirts	Water infrastructures are managed by a mix of communities, cooperatives, and households, and separate from other sectors [(a)] Outside the core residential areas (area of 2010), the communities have transferred the responsibilities to private consumers, who are also responsible for funding Specialized services are contracted to external companies Administrator leadership with focus on maintaining the Status Quo	As A3, but operation and maintenance is even worse, i e minimal	The infrastructures are decaying In the core residential areas (as 2010), water is centrally supplied, but it is not drinking water quality Households have own POU [(e)] systems to reach drinking water quality, or buy water in the supermarket The existing wastewater system is publicly operated In new urban areas, no pipe system is built Instead, storm water is infiltrated, or simply drained via roads and trenches and sanitary wastewater is decentrally treated with cheap technologies, e g septic tanks and a municipal collection service Household water in the outskirts is supplied by municipal trucks once a week
A5	Decaying infrastructure everywhere	Most infrastructure services as well as their funding are in the responsibility of the customers (households), who are well-informed The services are contracted to external organizations that have a long-term relationship with their customers	Measures are only undertaken upon urgent need for action; the replacement is in trench Operation and maintenance are minimal (as A4), and inspection is even worse, namely none at all	As in outskirts of A4: No centralized water supply, and no more pipes are built Storm water is infiltrated, or drained via roads and trenches The consumers operate tanks for drinking water that are recharged by private delivery service, and treat the water in house Household water is delivered by municipal trucks Wastewater is decentrally treated with cheap technologies, e g septic tanks and private collection service
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	Rehabilitation is done according to condition [(d)] Repair and replacement are done in trench Their operation, maintenance, and inspection are only moderate	The water supply and wastewater system are fully centralized, as in A1a, but with a much stronger focus on retention of storm water Water is supplied in drinking water quality, also for household use Water for

No.	Alternative name	Organizational structure, sector cooperation, management	Rehabilitation strategies, operation, and maintenance	Wastewater and water supply system technology
		telecommunication, electricity, gas, and road services [(a)] The constant budget is 100 % self-financed Management focuses on minimizing costs and maximizing performance, with strong personal motivation		firefighting is only partially supplied; further volumes are stored in underground tanks There is a 4th treatment step at the WWTP to remove micropollutants Separate systems and large stormwater retention volumes are installed in new development areas
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)], but there is no collaboration between different infrastructure services [(a)] Treatment facilities on private grounds (households, industry) are within the responsibility of the owner Management and funding as A6, but additionally well-informed households	Rehabilitation is done according to prioritization [(c)] Renovation is done in trench Their operation, maintenance, and inspection are only moderate (as A6)	The system is nearly fully decentralized Rainwater is reused in households as far as possible and treated at POE [(f)] Additional water will only be delivered with trucks by municipality upon special demand or in longer dry periods Water for firefighting is stored in shared community tanks (eggs) Wastewater is treated on-site, including urine source separation and nutrient recovery, with re-use as fertilizer in local agriculture The remaining wastewater is drained into the storm water sewers As in A6, large stormwater retention volumes will be installed in new development areas
A8a	Status Quo with storm water retention	The communities [(b)] remain responsible for a single, integrated wastewater and drinking water sector that jointly operate the water infrastructures, with some services contracted out to private enterprises Funding is flexible owing to a mix of 50 % leverage and 50 % self-finance Administrator leadership with focus on maintaining Status Quo	Rehabilitation is done according to prioritization [(c)] Renovation is trenchless Their operation, maintenance, and inspection is only moderate (as A6)	The water supply and wastewater system are fully centralized Drinking water is centrally supplied in large amounts, and can also be used for firefighting Storm water is infiltrated as much as possible Treatment of wastewater in the central WWTP as today
A8b–A8f	Status Quo technical variants	Organization of all variants as A8a	As A8a, except A8f where more demanding hydraulic design criteria are applied	The Status Quo is modeled with different technical variants of the current water supply and wastewater disposal system For example, there are separate or combined sewer systems; the system is extended or additional WWTP plants are built (A8b); decentral wastewater treatment with flush toilets (A8c); only one central WWTP for the whole region at different locations (A8d, A8e); water for firefighting is centrally distributed with drinking water, but newly developed housing areas have different dimensioned fire-flows (self-cleaning networks; A8b–A8f)
A9	Centralized, privatization, minimal maintenance	The water infrastructures are fully privatized, and all sectors work separately [(a)] Private consumers choose their contracting provider; in general all providers seek for revenue-maximization The constant budget is 0 % self-financed (100 % leverage) The well-informed households choose contractors, who have a long-term relationship with their customers	Measures are only undertaken upon urgent need for action; only repair is done, and in trench Operation and maintenance are minimal, with little inspection (as A4)	The water supply and wastewater system are fully centralized as in A8a Drinking water is centrally supplied in large amounts, but water for firefighting is only partially supplied Further volumes are stored in underground tanks Storm water is infiltrated as much as possible Treatment of wastewater in the central WWTP as today

[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

[c] x % of pipes in priority class y

[d] x % of pipes in condition y

[e] POU = Point of use treatment in the households to achieve drinking water quality; can be done e g on the tabletop or under the sink

[f] POE = Point of entry; e g water is treated where it enters the household water cycle at the entry point from a centralized water system or after a water storage tank

A4 Stakeholder feedback and recommendations

In all the first steps of Structured Decision Making (SDM, Gregory et al. 2012a) as developed in the SWIP-project and described above, we asked the stakeholders (interview partners or workshop participants) for feedback. We present details of this feedback in Table 36 below. We then summarize the advantages and disadvantages of the proposed approach, based on the stakeholder feedback. Finally, we give some recommendations for application the SWIP-approach for Structured Decision Making in other settings and applications.

Table 38: Stakeholder Feedback. Overview of the Steps of the SDM-process (Structured Decision Making), the types of stakeholder involvement and their feedback, the advantages and disadvantages of the adopted approach and recommendations for other applications. SH = stakeholders, WS = water supply, DW = drinking water, WW = wastewater, WWTP = wastewater treatment plant, CSOs = combined sewer overflows.

Step	Description of process	SH involvement	SH feedback	Advantages	Disadvantages	Recommendations
1	**Clarify decision context**					
1.1	**Case study selection and delimitation of system boundaries**					
	Intensive discussions to choose good case study; criteria: (a) good data ; (b) high pressure; (c) high motivation; (d) collaboration Detailed evaluation of 4 case studies Choice of "Mönchaltorfer Aa", mainly because of very high demand for collab-oration in NRP-61 (www nfp61 ch) Lack of strong pressure was later serious drawback: necessity to convince SH to participate	Phone, Email, meetings in case study area Clear definition of required data, time/ type of involvement (work-shops, interviews, questionnaires)	Resistance of some communities to participate Enablers: participation of other communities, acceptance by local politicians, national research programmed, good name of Eawag among engineers, support by engineering consultant Worries about time demands, type of involvement	Lengthy procedure resulted in good case-study knowledge and later high willingness of SH to collaborate Main advantage for research: sharing data with other projects in NRP-61 (agriculture, spatial planning of future scenarios, water quality)	Selection of case study clearly driven by request to collaborate and exchange data Problematic for MCDA, since local SH did not see need to change a system that is functioning well	Choose a real problem! i e SH need solution Clearly define interactions (type, number, length) Strong personal commitment of research team (e g organize attractive meetings) Look for support by important SH as mediators
1.2	**SH selection; clarify decision problem with SH**					
	Details in Lienert et al (2013) 1st stratified sampling: local, cantonal, national level (vertical axis), sectors (e g engineering, administration & politics; horizontal axis), all communities, WS & WW sector 2nd snowball sampling in 27 SH interviews Detailed feedback: who plays role in infrastructure planning, who is affected, interests, interactions Based on SH and network analysis: invitation to workshops; selection of 2 x 10 SH (WS/ WW) for MCDA preference elicitation interviews Feedback questions at end of interview (see 1 2 a - 1 2 e below)	Face-to-face interviews with 27 SH		Detailed SH and network analysis to select interview partners presumably only possible in research project Advantage: high confidence about very good representation of different interests In-depth knowledge about perspectives, current and future problems, interests, interactions, power relationships etc	Very time-consuming procedure; hardly feasible in real implementation projects with limited resources Not possible to cover representative population sample with face-to-face interviews	In most real applications, it might suffice to select SH with short questionnaire (Email, phone call, internet survey) Important questions: Who is involved in decision? Who is important to make decision on scale of 1 - 10? Who is affected once decision is made on scale of 1 - 10? What are their main interests? Who might have very different perspective?
1.2.a	**What is next step (by whom)?**					
	1st feedback question in 1st interview: what would be next step and who should do it? We categorized answers and state how often comments belonging to each category were made (in parentheses)	Face-to-face interviews with 27 SH	Eawag must do next step (mentioned 6x); include uncertainty in planning (6x); Eawag should show current state/ deficits (4x); increase professionalism of engineers (4x); performance of new	Clarification of expectation of SH: strong pressure on Eawag to do next step and concrete expectations about guidance in infrastructure planning, including	SH expect outcomes that surpass results of scientific project; may lead to disappointment if expectations are not satisfied As example, SWIP will not produce	Ask SH at early stage about expectations To avoid disappointment, clearly communicate type of results and which expectations can or cannot be satisfied (and why) We produced

Step	Description of process	SH involvement	SH feedback	Advantages	Disadvantages	Recommendations
			alternatives (3x); guidance w r t economic constraints (3x); planning tool (3x); strategy development (by authorities, professional organizations; 2x); support in planning/ enforcement of legislation (by canton; 2x) Mentioned 1x: training course at Eawag; better linked networks; information exchange between communities; end-users interests 10 of 27 respondents: no spontaneous idea	e g planning tool To a lesser extent: support in strategic planning and enforcement of legislation by authorities and organizations of water professionals	easy-to-use decision or planning tool as part of NRP-61 project; any such results will have to be pushed by project leaders at Eawag after termination of PhD-projects	information material specifically for communities and as preparation for interviews or workshops (see below)
1.2. b	**Expectations concerning Eawag?**					
	To clarify question 1 2 a, we then asked specifically for expectations, fears or hopes w r t our project and Eawag We categorized answers and state how often comments belonging to each category were made (in parentheses)	Face-to-face interviews with 27 SH	Eawag should generalize results, produce information material, guidelines, rationale to motivate communities to carry out strategic planning (9x); analysis of current situation (5x), basis for discussion in communities (4x); estimates about future of infra-structures (3x), networking (3x); decision tool (4x); analysis of non-conventional alternatives (1x); effects of micropollutants (1x); discuss results with authorities and national politicians (1x)	As above: ask about expectations of SH; e g with follow-up question to 1 2 a, as 1 2 b here	As above: risk of disappointing SH	As above: ask questions, ideally in different ways (1 2 a and 1 2 b); try to avoid disappointment and exaggerated expectations by clearly communicating expected results
1.2.c	**General feedback 1st interview: positive aspects**					
	Last question of 1st interview series: general feedback, separately for positive/ negative aspects/ recommendations We categorized answers and state how often comments belonging to each category were made (in parentheses) Here: positive aspects	Face-to-face interviews with 27 SH	Interview very agreeable/ good (13x); interview clear, well structured, well con-ducted, good questions (9x); no pos feed-back (8x); interesting/ important topic (5x); interview stimulated holistic thinking (4x); my view well acknowledged (2x); good science (1x); project carried out well (1x)	Clearly structure interview, carry it out in agreeable and respectful way	Possible disadvantage of strongly structured interview is restriction to specific questions, but this was not the case here since it was not criticized (see 1 2 d below) and interview even stimulated holistic thinking	Well prepare interview with clearly structured guideline, but leave room for creativity Treat respondents with respect, acknowledge their input, expertise and time (this should go without saying)
1.2. d	**General feedback 1st interview: negative aspects**					
	As above: general feedback, negative aspects	Face-to-face interviews with 27 SH	No neg comment (12x); not right expert (4x); interview cognitively very demanding (4x); too long (3x); focus not clear/ too local (3x); topic too abstract (climate change/ popup growth not relevant) (2x); questions not understandable for practitioners	Give room for negative feedback to better understand answers, improve further interactions and detect sensitive areas Problem about long and cognitively demanding interview can only be changed by reducing number and	Negative feedback may "hurt" Too late to change length of interview or type of questions to reduce cognitive effort at end of interviews	Give room for negative feedback Think strongly about language/ specific formulations to make questions understandable, avoid technical or scientific terms Consider trade-off between length of interview and required input Suggestion: general

Step	Description of process	SH involvement	SH feedback	Advantages	Disadvantages	Recommendations
			(complicated/ technical or scientific terms) (2x); discussion "ridiculous" (1x); fear of criticism of cantonal authorities (1x)	type of questions; must be decided within respective project		questions to all respondents; tailor sub-set to specific respondent to reduce time demand and cognitive effort
1.2.e	**General feedback 1st interview: recommendations**					
	As above: general feedback, recommendations	Face-to-face interviews with 27 SH	No recommendations (9x); interview other SH (5x); interest in results (7x); reduce time: strict guidance/ reduce questions (2x); reduce cognitive effort: simpler formulation/ reduce technical or scientific terms (2x) Mentioned 1x: send questions before; make questions more concrete; be respectful; keep survey anonymous; consider as many aspects as possible; analyze data neutrally; "just go and ask"	Recommendations are already reflected in positive and negative feedback above Answers from SH show that they come to similar conclusions as those drawn by ourselves, based on their feedback	See above	See above
2	**Define objectives and attributes**					
2.1	**Determine and discuss objective one-on-one with SH**					
	Preliminary objectives hierarchy set up by project team, based on engineering requirements for WS and WW system; additionally including "intergenerational equity", "high social acceptance" and "low costs" As 1st of interviews (see above), SH freely stated which objectives they need to compare alternatives We then showed and discussed our objectives Interviewee classified objectives into essential/ important/ nice to have We asked for attributes (details see Methods, main text)	Face-to-face interviews with 27 SH	Very detailed feedback concerning objectives, not presented here for reasons of space No feedback concerning method	Asking each SH alone and as first open question about objectives avoids priming effect and allows collecting their ideas, not ours By presenting and discussing our proposal: consolidation possible if they agreed that their objectives were covered in our hierarchy Further reduction by classification in "essential" obj (I really need these)	Face-to-face interviews is very lengthy procedure (for many SH) Risk of long list of objectives, important to only one or few SH Problem of how to reduce list if goal is concise objectives hierarchy that covers only most fundamental aspects SH are not aware of methodological requirements of objectives hierarchies	Trade-offs: avoid priming/ include many perspectives/ create short list of objectives Alternatively: create objectives in SH workshop (chance of "group opinion"; see below) To reduce objectives, directly pre-sent hierarchy; collect feedback w closed questions (e g : "do you agree"?) Process objectives thereafter to generate hierarchy applicable for all SH that meets method requirements (e g no redundancy)
2.2	**Discuss objectives hierarchy in 2nd SH workshop**					
	In 2nd SH workshop, we presented large objectives hierarchy with all "essential" objectives from interviews (see 2 1) and requirements for good objectives Participants systematically worked through hierarchy with neighbor; discussion which objectives are really needed or missing We collected notes and discussed objectives in plenum Each participant designated with points 3 least relevant objectives Presentation of our project, ourselves and especially MCDA-approach (see 3 1 below)	Workshop with 20 SH (identified with SH analysis, see 1 2 above)	Feedback concerning objectives see SI-A1 Feedback concerning workshop see 3 1 below	Ideally, workshop allows creating concise hierarchy that reflects all opinions Group process should allow better understanding of other SH (see e g Gregory et al 2012a) It should also be possible to agree on small number of really fundamental objectives that cover all important aspects	No deletion of objectives: no shared opinion w r t superfluous objectives "High social acceptance" and "inter-generational equity" most strongly questioned, but no clear justification for exclusion (plenary discussion/ points; see SI-A1, main text); they are also fundamental for sustainability Project team further processed hierarchy until it was complete, but as concise as possible (main text) Risk of missing fundamental objectives (early consensus)	Face-to-face discussion with neighbor gives voice to shy participants in workshop/ increases understanding Plenary discussions show breadth of opinions; ideally followed by process of focusing on fundamental objectives If not, consensus can be "forced" with moderation methods, e g assigning points But excluding objectives can be problematic; as in our case Workshop bears risk of losing control: good preparation/ moderation/ "emergency plans" Risk of missing out fundamental objectives ("groupthink" heuristic)
2.3	**Feedback to objectives and attributes during 2nd interview series (preference elicitation)**					

Step	Description of process	SH involvement	SH feedback	Advantages	Disadvantages	Recommendations
	We are currently carrying out face-to-face interviews with selected SH (see 1 2 above) to elicit their preferences for MCDA Elicitation of: scaling constants (weights), single-attribute value functions, aggregation scheme, risk attitude These interviews are not part of work presented here; but we give short overview of feedback concerning objectives and attributes, alternatives and MCDA procedure (for alternatives, see 3 2 below)	3 sets of face-to-face interviews (including reading information material and filling out online questionnaire before interview) with 10 SH in each set SH identified with SH analysis (see 1 2 above)	Understandability Difficulty to understand: some objectives (e g "good chemico-physical quality" of DW)/ highly uncertain attributes (e g "few gastrointestinal infections")/ attributes w complex models (e g "good chemical state of watercourse": 1 classify pollutants, 2 mathem aggregate each reference point in catchment, 3 spatially aggregate) Missing or irrelevant objectives e g "technology readiness"; "high redundancy of WS"; replace "good chemical state" & "low neg hydraulic impacts" w "good eco-morphol state" Attribute ranges (see SI-A1) Some doubts; e g 100 % co-determination of end-users (is not desirable); worst cases for Switzerland of "high reliabil of drainage system"; "good chemic state of watercourse" (unrealistic); worst case of 60 CSOs ("few gastro-intest infect "; too optimistic) Preferential independence Some objectives not preferentially independent for all SH, e g "low future rehabilitation burden" & "exfiltration from sewers"/ "intergenerational equity" & "low costs"/ "high reliability of WS" & "good quality of DW" Minimum criteria In 7 of 10 WS interviews: alternatives not fulfilling minimal water quality standards are not acceptable e g "microbial and hygienic quality" of DW For WW, few SH regard current laws as minimal standards/ or as optimal level	Understandability Wherever possible, we used attributes common in the field (e g chemical state of water) Hence, they are backed by natural-scientific evidence/ we can rely on real numbers (expert estimates or models) Missing or irrelevant objectives Irrelevant objectives can easily be dealt with by giving zero weight Ranges Where possible, we defined attributes/ ranges to be applicable to other case studies/ to "Boom" or "Doom" scenario Preferential independence If this requirement holds, the simple linear additive model can be used Minimum criteria Minimal requirements can easily be implemented in MCDA, i e with MAVT/ MAUT by aggregating values from lower to higher levels of objectives hierarchy such that higher-level value is never better than worst value achieved for lower-level objective Or exclude all alternatives that do not fulfill minimum requirement	Understandability Relatively technical or natural-scientific attributes are not necessarily easy to understand for non-experts Missing or irrelevant objectives Missing objectives cannot be added at later stage of MCDA (respective SH has to accept this) Follow-up option: evaluate sensitivity of best-performing alternatives of MCDA w r t missing objective Ranges Attribute levels of broad ranges might seem unrealistic Can make preference elicitation more difficult and affects weights if ranges are not adequately considered Preferential independence If this requirement does not hold, more complex aggregation models (e g multiplicative) are needed (elicit additional scaling constant!) Minimum criteria Disadvantage of minimal requirements: all alternatives that do not fulfill requirements are excluded or receive equal overall values/ utilities (but alternative that performs better regarding other objectives should probably receive higher values)	Understandability Elicit single-attribute value functions for technical/ natural-scientific attributes from experts (e g "good chemical state") Use expert value functions for other SH/ elicit only scaling constants (weights) Missing or irrelevant objectives Create objectives hierarchy carefully/ with intensive SH interaction (as our example) If SH doubts result because of missing objective, carry out rough estimate of sensitivity of best-performing alternatives to this objective Exclude irrelevant objectives with zero weight Ranges Define attributes to be generalizable, allow-ing for up- or down-scaling in other cases; e g use relative, not absolute numbers ("number of pipe failures/ yr/ 1'000 inhabitants" instead of "number of pipe failures") To make relative numbers tangible, pre-sent absolute numbers for case study (see "Status Quo"; Online Res 7) If SH think ranges are unrealistic, use example of countries w lower infrastructure standards During elicitation, point out ranges repeatedly to avoid "range effect" bias Preferential independence Construct objectives hierarchy to fulfill this, but check validity in interview (e g "do preferences about one attribute depend on level of another?") Minimum criteria Discuss implication of minimal requirements (see disadvantages) with SH
	Develop future scenarios					
	1st SH workshop					
	Scenario planning not included in standard MCDA Aim: capture future uncertainty w r t socio-economic development with snap-shot images We invited 22 community members, excluding high-rank officers to create good workshop feeling 15 SH participated (all 4	Workshop with 15 local SH (identified with SH analysis, see 1 2 above)	Summary "what would I be happy about": * know other participants, teambuilding * know region, networks for communities * identification with project	We used 1st workshop to introduce SWIP-project; good opportunity to get local participants "on board" Positive (less intimidating) to only include local SH Scenario planning approach was clearly	Thinking in extreme scenarios might create impression that we are not dealing with the real problems of SH Scenario workshop needs to be very well prepared and moderated: convey	Use workshops to introduce project and scientists Construct groups to later profit from "group feeling"; e g concerning collaboration across communities Whom to include or exclude? e g invite only local SH (as in

Step	Description of process	SH involvement	SH feedback	Advantages	Disadvantages	Recommendations
	communities, both water sectors, different roles) We first presented ourselves and SWIP Scenarios set in year 2050, discussed/ adapted to local case in 3 groups (equal distribution of perspectives) We used 4 Swiss scenarios from Swiss National Research Programme (NRP 54; www nfp54 ch) as framework Specification to local case based on variation of 8 factors, relevant for water infrastructures Scenarios visualized, presented, discussed in plenum (see main text, SI-A2) Feedback: "what would I be happy about?"/ "what learning effect did I have?"		* concrete results/ tool cost-benefit calcul * better understand objectives and output * deal w unpredictable future scenarios * should be exciting <u>Summary of "learning effect"</u>: * good discussion, excellent group work * fruitful/ creative method * good to think of future * surprised about scenarios/ not realistic * necessary to consider extreme scenarios * challenge to deal w results in real world * good exchange/ collaborat w commun * lots of fun/ creative, now back to reality	highly stimulating, very creative and lots of fun It helped to create team-feeling; raise interest in project Scenario planning invites thinking in broader terms about future in region, than what is usually done	that it is real science, despite being fun Only limited participants in scenario workshop; else discussion is likely less productive	our example)? Only limited number of participants in a workshop: how to select most important ones (e g SH analysis)? Very careful preparation and moderation, since things easily get out of control when "playing around" Decide about using framework (as we did), or creating scenarios from scratch
3	**Identify and create decision alternatives**					
3.1	**2nd SH workshop to identify decision alternatives with help of a strategy generation table and the future scenarios**					
	First, we introduced project and MCDA-approach, and discussed objectives (see 2 2 above) Creation of alternatives in 2nd SH workshop To stimulate creativity, we used 4 socio-economic scenarios as background (see above) We prepared “strategy generation table” (Howard 1988; Gregory et al 2012b): 17 basic factors (organizational structure, geographic extent, financial strategy, construction & operation of infrastructure, WW and DW system technology) Each factor has a number of specifications; a decision alternative consists of plausible combinations, which were created in workshop 20 workshop participants split into 4 mixed groups, each assigned to a scenario Each group created at least 2 strategic alternatives by choosing plausible specification for each factor (SI-A3) Project team later processed and detailed the 10 strategic decision alternatives from workshop	Workshop with 20 SH (identified with SH analysis, see 1 2 above)	No systematic collection of written feed-back to this workshop; following is based on our own impression We did ask SH to indicate on a poster-sized x-y-grid "how pleased are you with workshop?" (x-axis) and "how confident are you about SWIP-project?" (y-axis; scale "very low" to "very high") 10 of 20 participants gave feed-back High satisfaction with workshop (all above medium), but fairly low confidence in project: 3 points below medium and 3 exactly on medium line (others above)	Due to participation, SH understands methods; alternatives are relevant to SH; increases later acceptance of results; avoids overlooking issues obvious to local practitioners Combining“ strategy generation table” with scenarios is highly effective to avoid anchoring on Status Quo alternatives (Nutt 2004) e g generation of conventional central WS & WW treat-ment alternatives under “Status Quo”; “Boom” scenario triggered high-tech on-site solutions; “Doom” scenario cheap/ simple alternatives SI-A3) Characteristics of good alternatives: complete, com-parable, value-focused (addressing what matters), fully specified, intern-ally coherent, distinct (e g Gregory et al 2012a; Keeney and Raiffa 1976) This is well addressed by strategy generation table: SH are forced to rigorously cover important elements; increases	Clear disadvantage of strategy generation table: not very creative; not much fun; choosing specification for each factor is tedious work We later had to commit considerable work to further specify alternatives and include new factors missing in first version of strategy generation table, but important to distinguish alternatives Strategy generation table is rather time consuming Duration of workshop was about 6 hours (3 hrs for objectives (see 2 2 above); 3 hrs creating alternatives) SH were tired at end of day and strategy generation was done under time pressure We think that negative feedback concerning "confidence about project" might have been caused by this fatigue, possibly combined with some doubts about MCDA-approach, which seems somewhat difficult to understand	To reduce feeling of boring work, we recommend creating storylines about strategic decision alternatives with SH in workshop Carry out factor specifications later by project team w strategy generation table Combination ensures that SH are involved, i e that alternatives are adapted to local needs, make use of their know-ledge and are later better accepted, but that they do not lose interest Apart from "not much fun" aspect, we find strategy generation table a highly useful and systematic approach that ensures coverage of different aspects/ internal consistency We recommend combining a rigorous approach (e g strategy generation table) with very creative approach (e g scenarios as background) to avoid anchor-ing effects and focus on Status Quo Make sure to assign ample time Because MCDA-approach seems difficult to under-stand (general feedback that we receive again and again), we recommend to use every opportunity to present method; e g as

Step	Description of process	SH involvement	SH feedback	Advantages	Disadvantages	Recommendations
				internal consistency (SI-A3)		introduction to workshop
3.2	**Feedback to alternatives during 2nd interview series (preference elicitation)**					
	We are currently carrying out face-to-face interviews with selected SH (see 1 2 above) to elicit their preferences for MCDA Elicitation of: scaling constants (weights), single-attribute value functions, aggregation scheme, risk attitude These interviews are not part of work presented here; but we give short overview of feedback concerning the alternatives (for the objectives, attributes and general feedback, see 2 3 above)	3 sets of face-to-face interviews (including reading information material and filling out online questionnaire before interview) with 10 SH in each set SH identified with SH analysis (see 1 2 above)	Understandability of hypothetical alternatives One SH had difficulty to evaluate hypothetical alternatives that are very different today (feedback w r t separate supply of DW/ water for household/ for firefighting) Some hypothetical alternatives are un-realistic, e g one SH found it impossible to imagine a system which realizes all rehabilitation demand but has very low reliability Comparability of hypothetical alternatives Two SH found costs of WS (5 % of average annual income) as totally unrealistic for Switzerland ("American circumstances"!), thus difficulty to answer trade-off questions for hypothetical alternatives using this attribute level Trade-off questions difficult if they invoke moral conflicts, e g trade-offs between "few gastrointestinal infections through contact with WW" and "good chemical state of watercourse" Some trade-off questions ask respondents to choose between two unsatisfactory alter-natives, which gives uncomfortable feeling (Methodical issues, e g concerning elicitation with trade-off will be addressed in more detail in later papers)	Understandability of hypothetical alternatives To broaden range of decision alternatives also unconventional (but existing) solutions should be considered In current Switzerland, decentralized, on-site solutions and solutions (e g addressing water scarcity) are rarely discussed, but may become more viable in the future (climate change) and are certainly under discussion in more arid regions (e g Australia)	Understandability of hypothetical alternatives We included uncommon, somewhat visionary decision alternatives, which seem difficult to assess for some SH A remaining methodological problem is the construction of hypothetical alternatives that result in unrealistic combinations Comparability of hypothetical alternatives It is problematic that we had to set the ranges so broadly (see 2 3 above), which results in having to compare hypothetical and extreme alternatives Generally, trade-off questions seem difficult to answer, especially if they invoke moral conflicts and/ or leave the respondent feeling uneasy about his or her choice	Understandability of hypothetical alternatives SH should be included in generating decision alternatives (see 3 1 above), to make more exotic decision alternatives better tangible Comparability of hypothetical alternatives Reasons for working with extreme or unconventional alternatives and broad ranges must be explained as well as possible to SH We wish MCDA-procedure to be well-applicable to other cases, and to hold under different future scenarios However, problem remains that some methods force respondents to make morally difficult choices We discuss this in later papers about elicitation methods Current recommendation: choose elicitation methods that do not require extreme hypothetical alternatives/ not very difficult moral choices

SI-B) Supporting Material to: Combining expert knowledge and local data for improved service life modeling of water supply networks

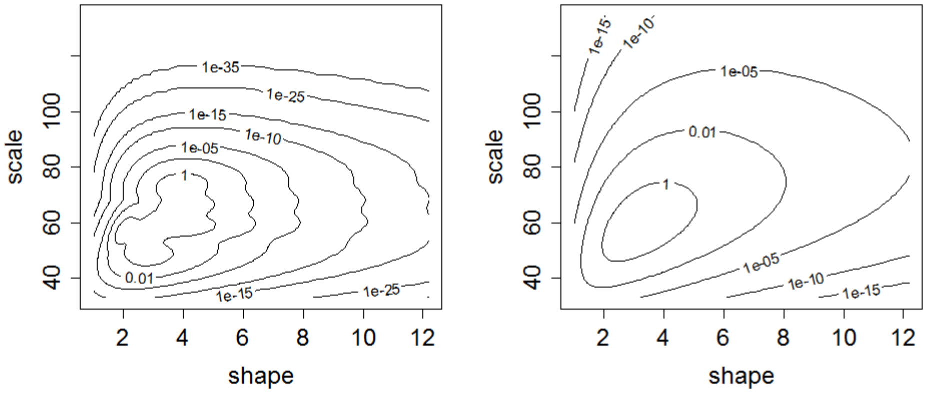


Figure B.1: Bivariate probability density distribution of the aggregated prior before smoothing (left, multimodal) and after smoothing (right, unimodal) for ductile cast iron (1965-1980).

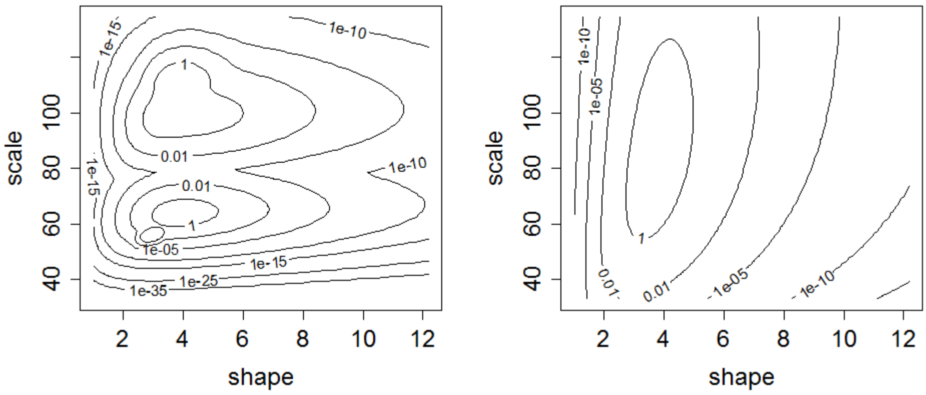


Figure B.2: Bivariate probability density distribution of the aggregated prior before smoothing (left, multimodal) and after smoothing (right, unimodal) for asbestos cement.

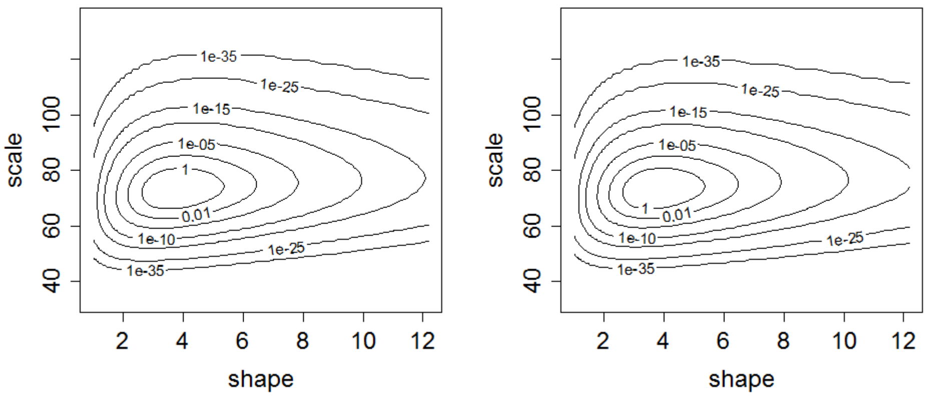


Figure B.3: Bivariate probability density distribution of the aggregated prior before smoothing (left, multimodal) and after smoothing (right, unimodal) for steel.

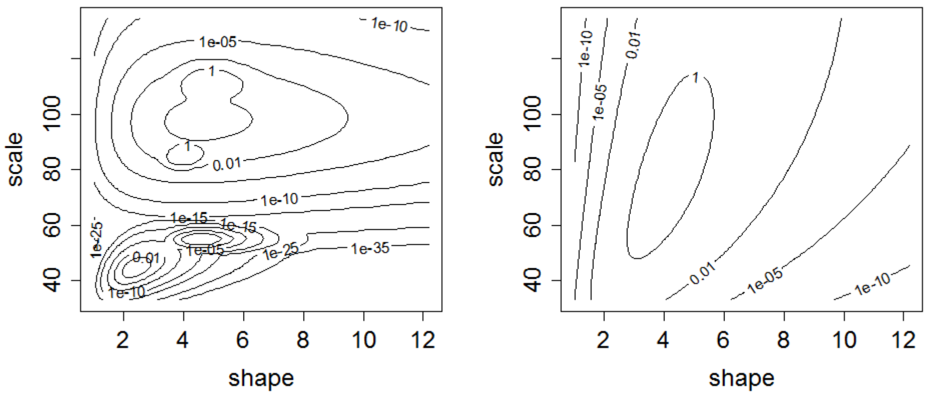


Figure B.4: Bivariate probability density distribution of the aggregated prior before smoothing (left, multimodal) and after smoothing (right, unimodal) for polyethylene.

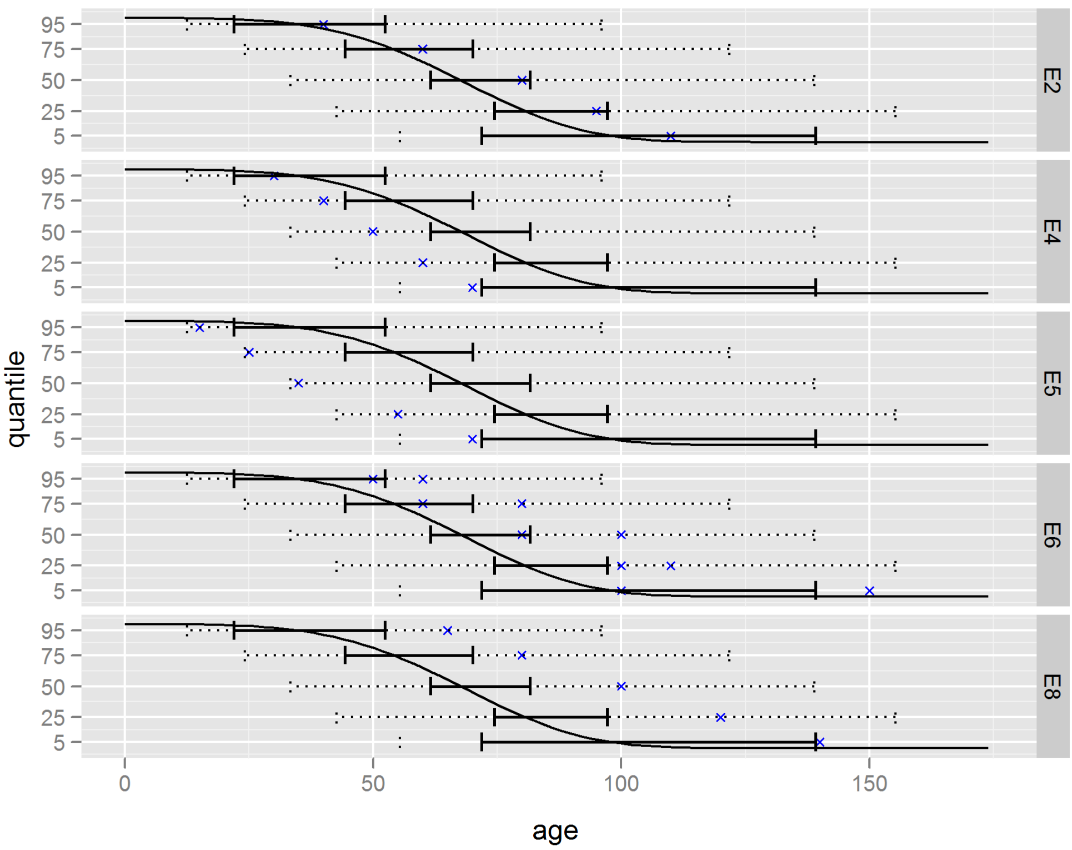


Figure B.5: Comparison of ST priors and estimates from experts. Blue crosses represent quantile values as stated by the expert indicated on the right edge (E1…E8). Solid error bars give the 95 % confidence intervals for complete pooling, dotted error bars for partial pooling. The survival curve is calculated from the mean parameters $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ of the partial pooling prior, see Table 8.

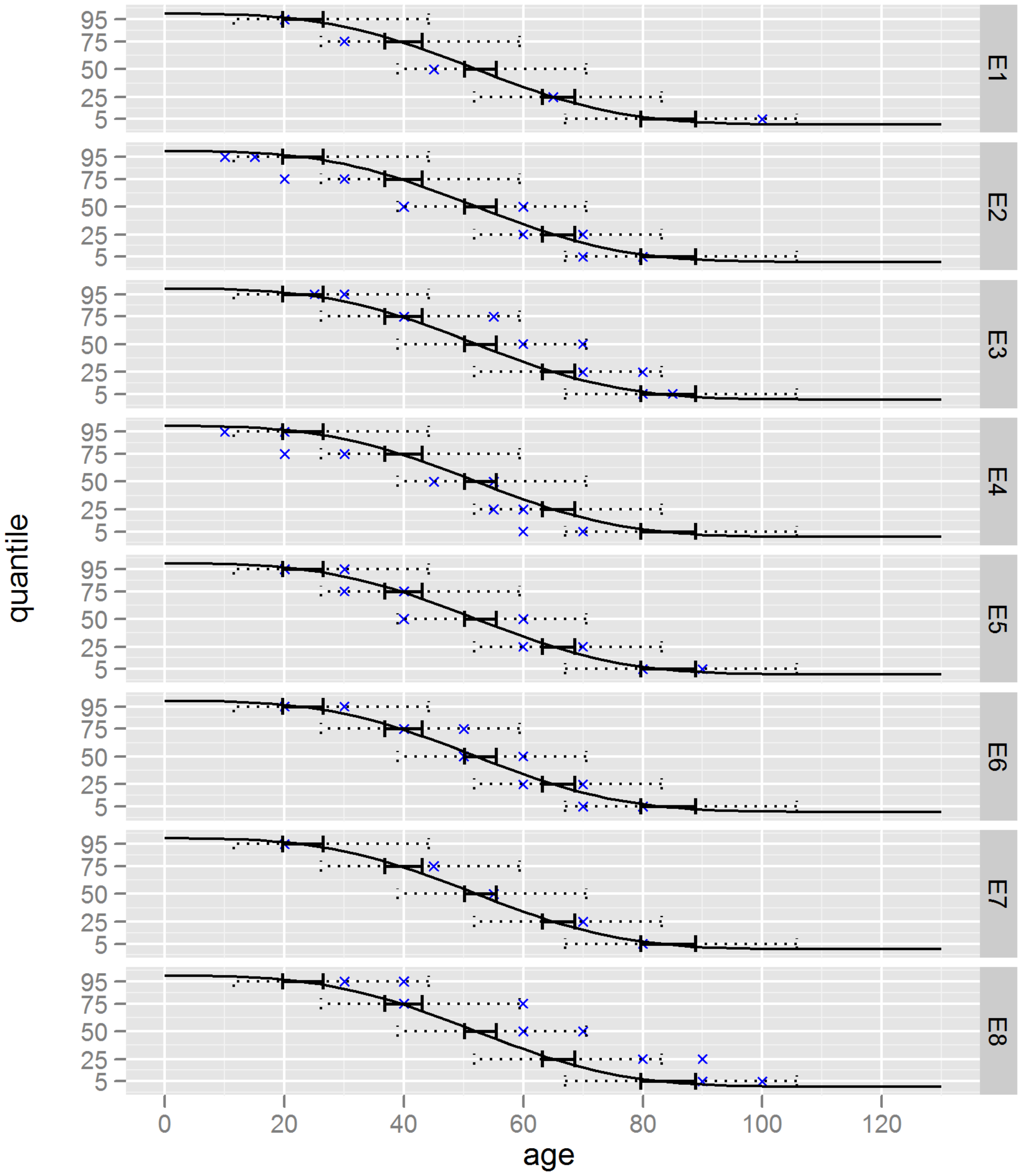


Figure B.6: Comparison of DI1 priors and estimates from experts. Blue crosses represent quantile values as stated by the expert indicated on the right edge (E1…E8). Solid error bars give the 95 % confidence intervals for complete pooling, dotted error bars for partial pooling. The survival curve is calculated from the mean parameters $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}})$ of the partial pooling prior, see Table 8.

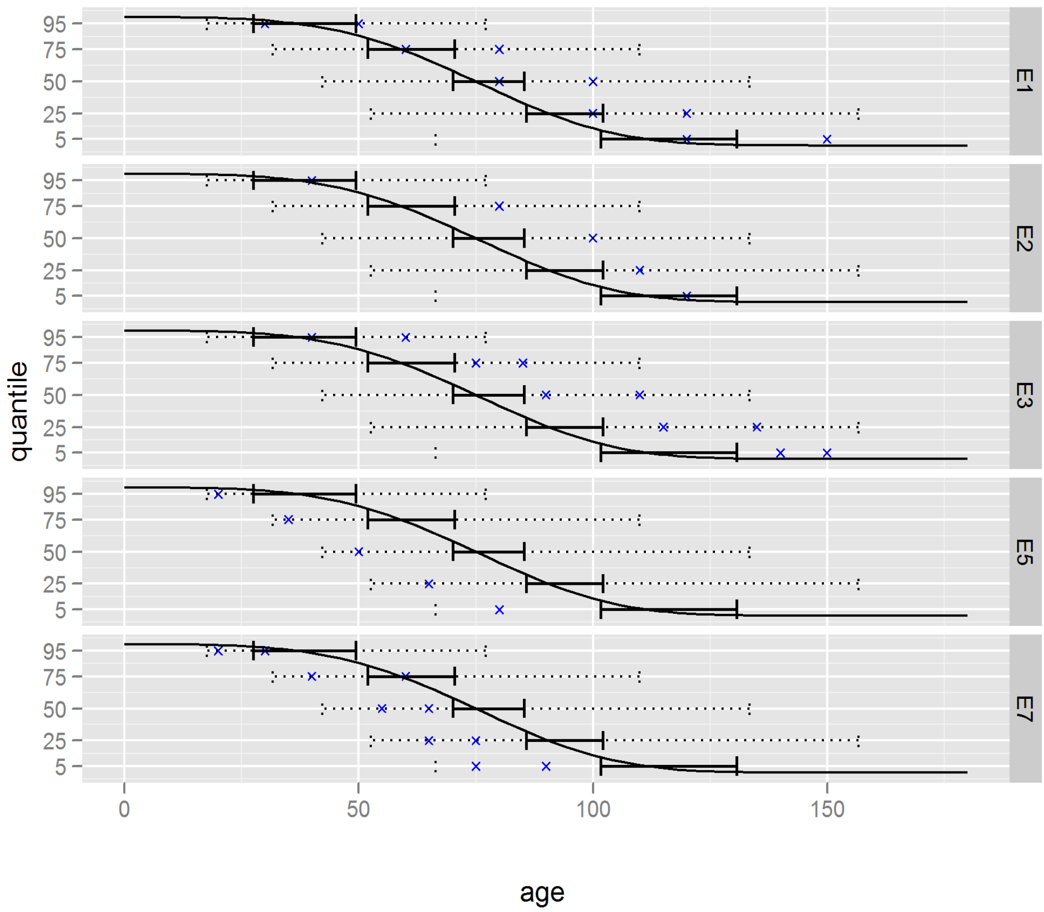


Figure B.7: Comparison of AC priors and estimates from experts. Blue crosses represent quantile values as stated by the expert indicated on the right edge (E1…E8). Solid error bars give the 95 % confidence intervals for complete pooling, dotted error bars for partial pooling. The survival curve is calculated from the mean parameters $(\hat{\alpha}, \hat{\beta})$ of the partial pooling prior, see Table 8.

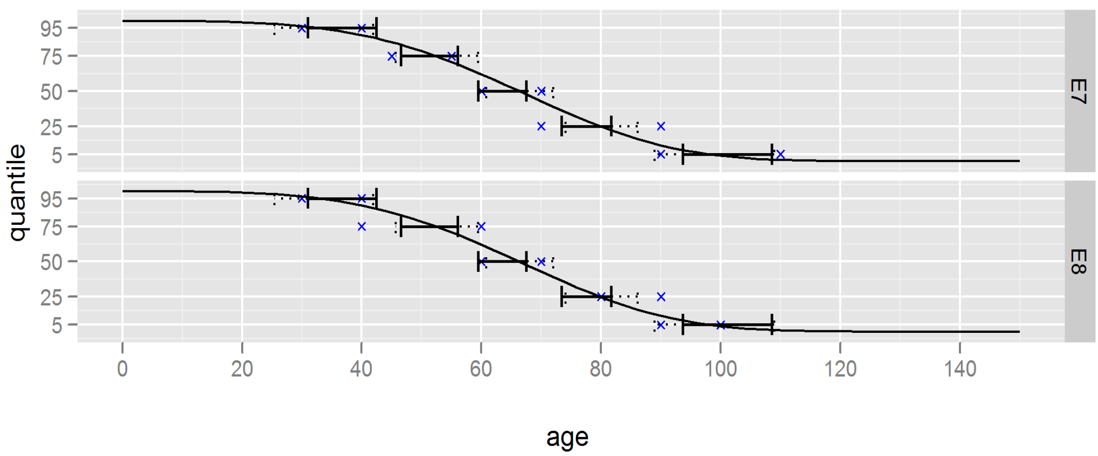


Figure B.8: Comparison of ST priors and estimates from experts. Blue crosses represent quantile values as stated by the expert indicated on the right edge (E1…E8). Solid error bars give the 95 % confidence intervals for complete pooling, dotted error bars for partial pooling. The survival curve is calculated from the mean parameters $(\hat{\alpha}, \hat{\beta})$ of the partial pooling prior, see Table 8.

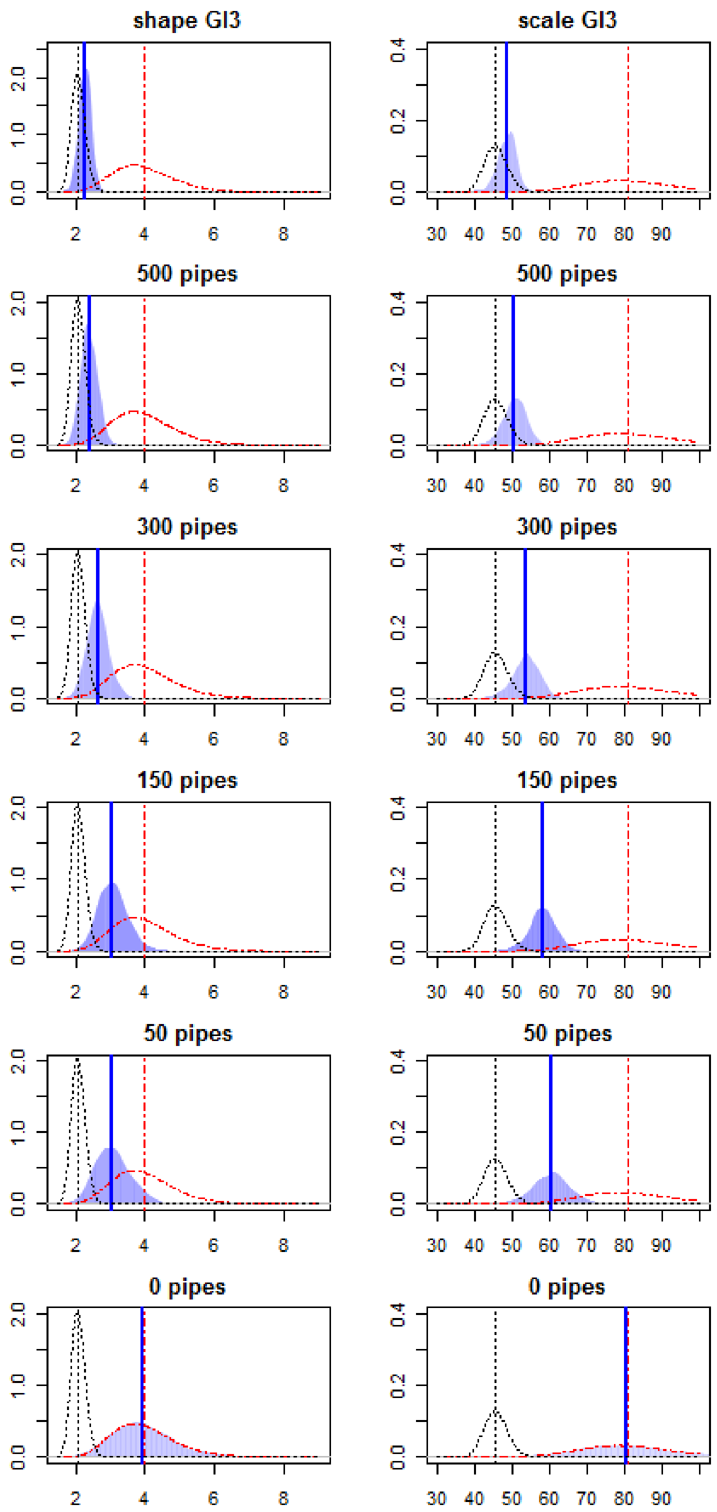


Figure B.9: Bayesian inference with partial-pooling prior for grey cast iron (GI3): Posterior (blue filled), and prior (red dash-dotted) marginal distributions of the Weibull shape (left column) and scale (right column) parameters for varying amounts of data (top level = all data). As a reference, the distributions resulting from MLE with all data (black dotted) are also plotted. Vertical lines indicate the position of the corresponding means.

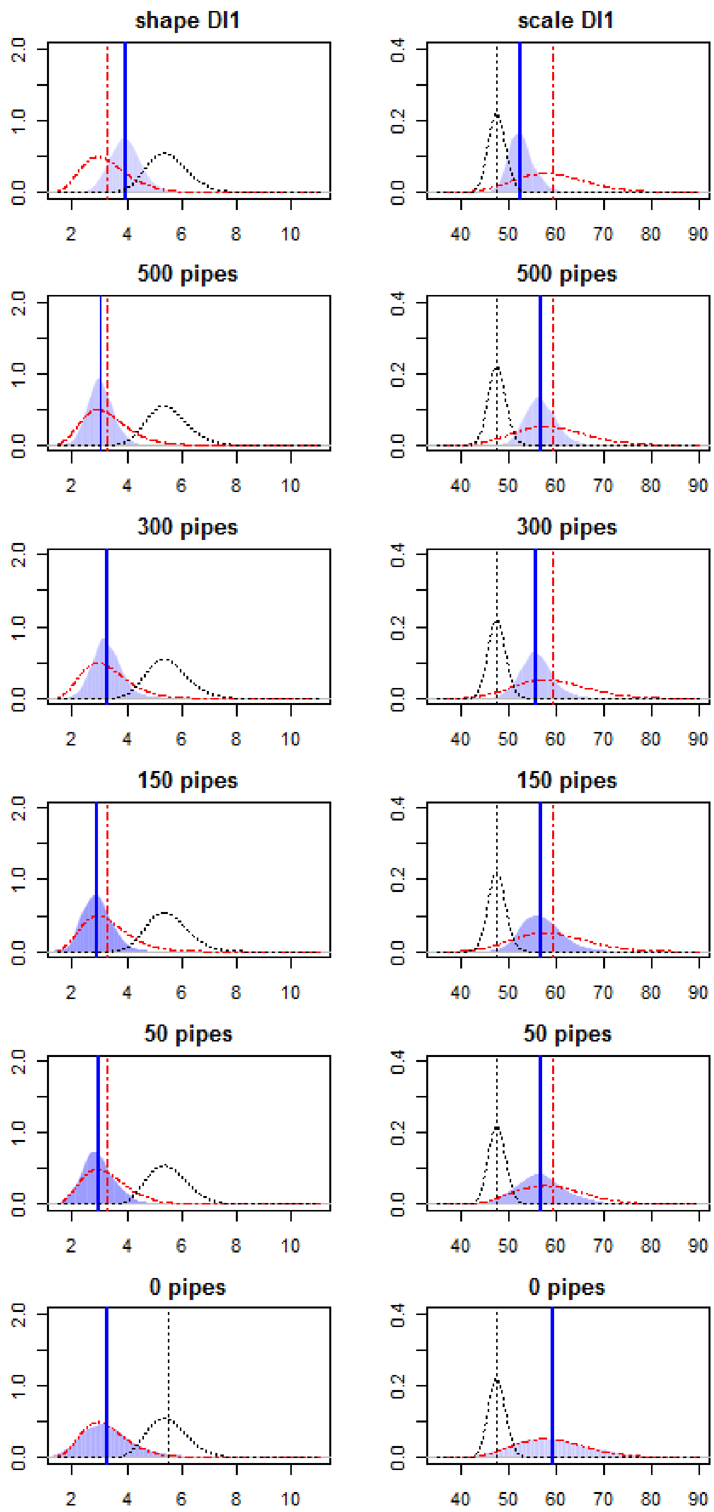


Figure B.10: Bayesian inference with partial-pooling prior for ductile cast iron (DI1): Posterior (blue filled), and prior (red dash-dotted) marginal distributions of the Weibull shape (left column) and scale (right column) parameters for varying amounts of data (top level = all data). As a reference, the distributions resulting from MLE with all data (black dotted) are also plotted. Vertical lines indicate the position of the corresponding means.

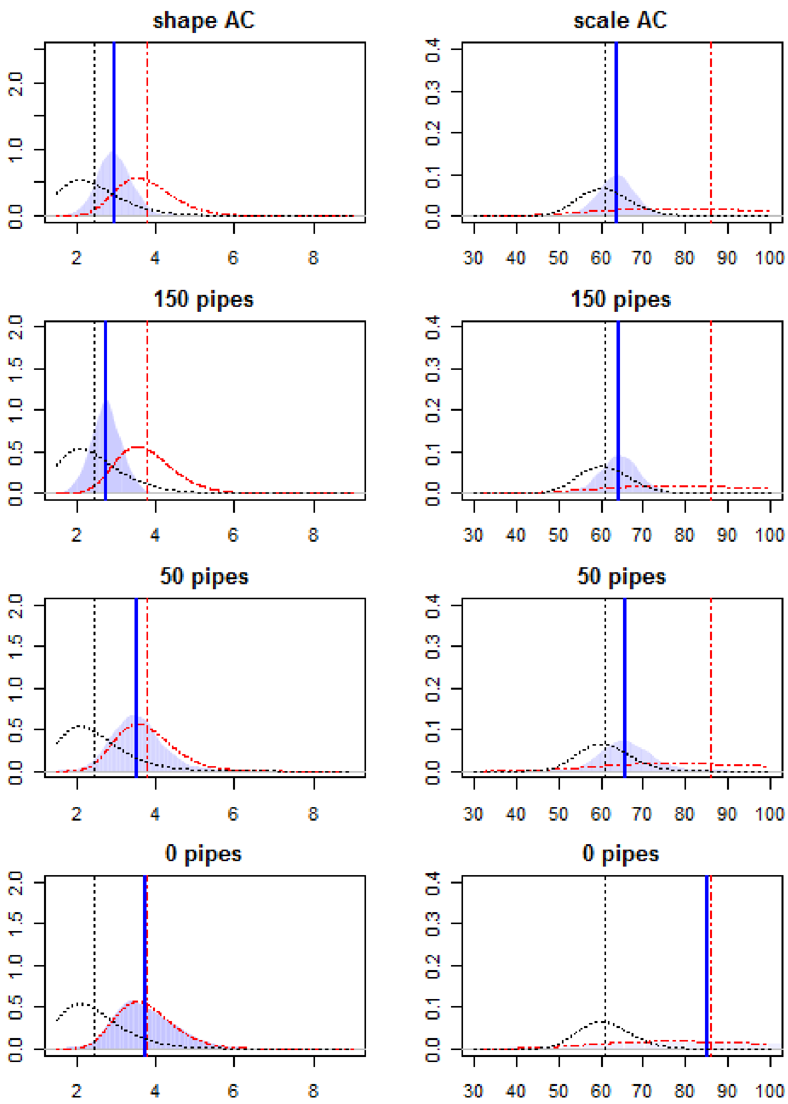


Figure B.11: Bayesian inference with partial-pooling prior for asbestos cement: Posterior (blue filled), and prior (red dash-dotted) marginal distributions of the Weibull shape (left column) and scale (right column) parameters for varying amounts of data (top level = all data). As a reference, the distributions resulting from MLE with all data (black dotted) are also plotted. Vertical lines indicate the position of the corresponding means.

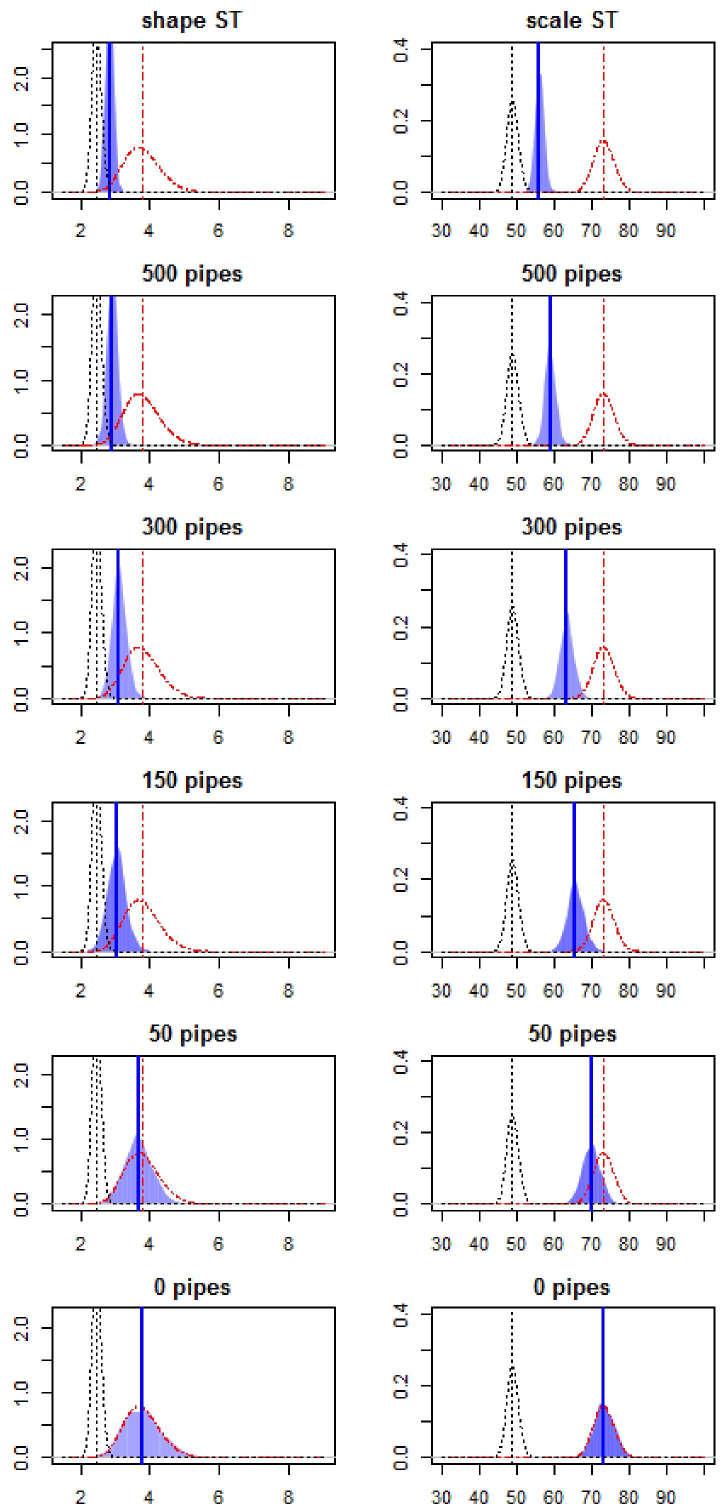


Figure B.12: Bayesian inference with partial-pooling prior for steel: Posterior (blue filled), and prior (red dash-dotted) marginal distributions of the Weibull shape (left column) and scale (right column) parameters for varying amounts of data (top level = all data). As a reference, the distributions resulting from MLE with all data (black dotted) are also plotted. Vertical lines indicate the position of the corresponding means.

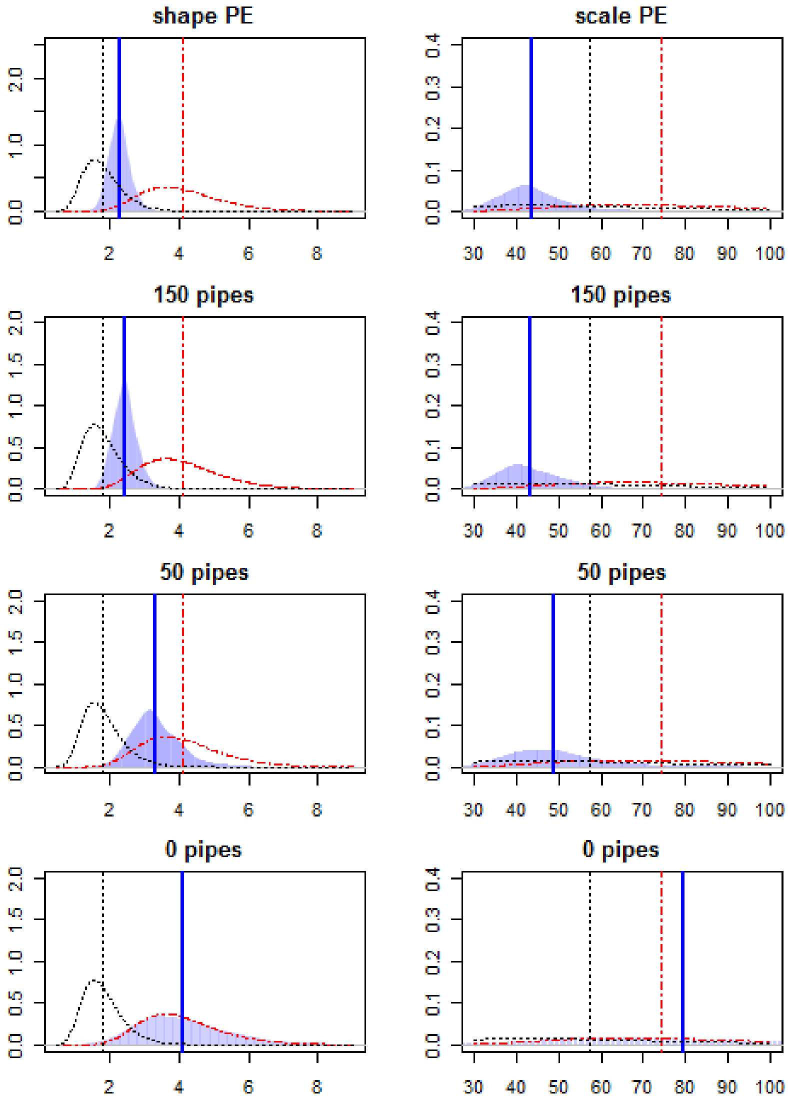


Figure B.13: Bayesian inference with partial-pooling prior for polyethylene: Posterior (blue filled), and prior (red dash-dotted) marginal distributions of the Weibull shape (left column) and scale (right column) parameters for varying amounts of data (top level = all data). As a reference, the distributions resulting from MLE with all data (black dotted) are also plotted. Vertical lines indicate the position of the corresponding means.

SI-C) Supporting Material to: Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis

C1 Symbols and abbreviations

SYMBOL OR ABBREVIATION	INTERPRETATION
Main text	
A,B	Hypothetical alternatives A, B
$A_{a0.5\%...2\%}$	See explanation in Tab.1 of the main text.
$A_{cyc80...100}$	See explanation in Tab.1 of the main text.
A-D	Four water utilities: A, B, C, and D
$A_{f0.5\%...2\%}$	See explanation in Tab.1 of the main text.
$A_{f2...5+}$	See explanation in Tab.1 of the main text.
$A_{fr1\%...2\%}$	See explanation in Tab.1 of the main text.
A_{ref}	See explanation in Tab.1 of the main text.
β_m	Regression coefficient / covariate
C_i	Criticality index (importance weight) of pipe diameter group i
c_j	Constant that determines the curvature of marginal value function over the attribute linked to objective j.
DI1	First generation ductile iron; centrifugal casting, before 1980
DI2	Second generation ductile iron; centrifugal casting, after 1980
D_{reha}	Degree of rehabilitation
ELECTRE	ELimination Et Choix Traduisant la REalité (Elimination and Choice Expressing Reality)
$P_A(X_A)$	Cumulative distribution of X_A
FAST	Fichtner asset services and technologies (asset management software)
FC	Fiber cement/asbestos cement incl. Eternit
GI2	Second generation grey cast iron; vertical casting, before 1930
GI3	Third generation grey cast iron; centrifugal casting, after 1930
IAM	Infrastructure asset management
k	Pipe index
m	Pipe characteristic, e.g. material
μ'_A, μ'_B	Mean of alternative A, B
μ'_A, μ'_B	Risk-adjusted mean of alternative A, B
μ_u	Parameter vector of means of the multivariate normal distribution
MAUT	Multi-attribute utility theory
MAVM	Multi-attribute value model
MAVT	Multi-attribute value theory
MCDA	Multi-criteria decision analysis
$n_{f,i}$	Number of pipe failures in pipe diameter group i
n_i	Number of pipes in pipe diameter group i
PE	Polyethylene
R	system reliability
r_{ref}	Failure rate of the reference strategy A_{ref} [#/(km*a)]
r_s	Failure rate of strategic alternatives s [#/(km*a)]
$\sum_u$	Parameter vector of standard deviations of the multivariate normal distribution
SAM	Strategic asset management
ST	Steel

SYMBOL OR ABBREVIATION	**INTERPRETATION**
t	Evaluation year
t_0	Laying year
V(*A*)	Aggregate value of alternative *A*
v.1cc.eqw, v.2cc.eqw, v.3cc.eqw	See Tab.2 in main text
v.1cv.eqw, v.2cv.eqw, v.3cv.eqw	See Tab.2 in main text
v.acv.eqw, v.acc.eqw	See Tab.2 in main text
v.lin.eqw	See Tab.2 in main text
v.lin.w1a, v.lin.w2a, v.lin.w3a	See Tab.2 in main text
v.lin.w1h, v.lin.w2h, v.lin.w3h	See Tab.2 in main text
$v_j(x_j)$	(Marginal) value function over the attribute linked to objective j
$v_j(x_j(A))$	(Marginal) value function over attribute linked to objective j of alternative *A*
w_1, w_2, w_3	See Tab.2 in main text
w_j	Weight of objective j
X_A	Random variable describing the attribute outcome of alternative A
x_j	Attribute level regarding objective j
$z_{k,j}$	Indicator variable, equals 1 if *j*th characteristic is met, else 0.
θ	Failure model parameter vector
θ_1	Weibull shape parameter
θ_2	Weibull scale parameter
θ_3	Exponential scale parameter
π	Probability not to be replaced after a failure
Appendices	
cr	Scenario-dependent change rate
g_1,g_2	Adjustment factors to account for changing diameter proportions in the overall pipe network
l_p	Future per person expansion length
$l_{p,0}$	Current per person expansion length
P	Population
P_0	Original population in reference year T_0
T	Evaluation year; here= 2010
T_0	Reference year; here = 2010

C2 Prediction of unrecorded failures

The number of failures of a pipe between its date of laying t_o and the beginning of the failure recording period *a* is distributed according to

$$Prob\left(n^{(1)}|n^{(2)}, t_1^{(2)} \dots t_{n^{(2)}}^{(2)}, in\ [t_o, a]\right) = \int_{t_0}^{t_1^{(1)}} \dots \int_{t_{n^{(1)}-1}^{(1)}}^{a} \frac{p\left(n^{(1)} + n^{(2)}, t_1^{(1)} \dots t_{n^{(1)}}^{(1)}, t_1^{(2)} \dots t_{n^{(2)}}^{(2)} | in\ [t_o, b]\right)}{p\left(n^{(2)}, t_1^{(2)} \dots t_{n^{(2)}}^{(2)} | in\ [a, b]\right)} dt_{n^{(1)}}^{(1)} \dots dt_1^{(1)} \tag{C.1}$$

The distribution is conditioned on the known $n^{(2)}$ observed failures at $t_1^{(2)} \dots t_{n^{(2)}}^{(2)}$ within the observation period $[a, b]$. The enumerator is given in equations (14) and (15) in (Scheidegger et al., 2013), or (26) and (27) in chapter 4. To sample from (C.1) an expression that is proportional to it is sufficient so the evaluation of the denominator is not required.

C3 Estimated failure model parameters for runs with fixed π in water utilities B and C

Table C.1: Summary statistics of parameters after inference with fixed π in water network B and C

		B						C					
	π [quantile]	$\hat{\theta}_1$	$\hat{\theta}_{2,DI1}$	$\hat{\theta}_{3,DI1}$	$\hat{\beta}_{DI2}$	$\hat{\beta}_{GI3}$	$\hat{\beta}_{2,FC}$	$\hat{\theta}_1$	$\hat{\theta}_{2,DI1}$	$\hat{\theta}_{3,DI1}$	$\hat{\beta}_{DI2}$	$\hat{\beta}_{GI3}$	$\hat{\beta}_{2,FC}$
$\hat{\boldsymbol{\theta}}$	**0.619** [0.01]	**1.69**	**45.65**	**15.50**	**2.95**	**1.28**	**5.25**	**1.28**	**64.73**	**11.26**	**3.26**	**0.91**	-
$sd(\hat{\theta})$		*0.27*	*3.35*	*3.20*	*0.93*	*0.14*	*1.63*	*0.17*	*6.53*	*1.58*	*1.06*	*0.13*	-
$\hat{\boldsymbol{\theta}}$	**0.745** [0.1]	**1.71**	**52.55**	**18.50**	**2.83**	**1.28**	**5.05**	**1.36**	**81.44**	**13.60**	**2.73**	**0.96**	-
$sd(\hat{\theta})$		*0.27*	*4.09*	*3.78*	*0.84*	*0.13*	*1.51*	*0.18*	*8.90*	*1.88*	*0.79*	*0.11*	-
$\hat{\boldsymbol{\theta}}$	**0.793** [0.2]	**1.72**	**55.00**	**19.65**	**2.79**	**1.28**	**5.00**	**1.38**	**87.07**	**14.57**	**2.60**	**0.97**	-
$sd(\hat{\theta})$		*0.27*	*4.48*	*4.00*	*0.82*	*0.13*	*1.48*	*0.18*	*9.99*	*2.01*	*0.72*	*0.11*	-
$\hat{\boldsymbol{\theta}}$	**0.825** [0.3]	**1.72**	**56.59**	**20.41**	**2.77**	**1.28**	**4.96**	**1.40**	**90.59**	**15.23**	**2.52**	**0.98**	-
$sd(\hat{\theta})$		*0.27*	*4.76*	*4.14*	*0.81*	*0.13*	*1.46*	*0.18*	*10.74*	*2.09*	*0.69*	*0.11*	-
$\hat{\boldsymbol{\theta}}$	**0.850** [0.4]	**1.73**	**57.82**	**21.01**	**2.75**	**1.28**	**4.93**	**1.41**	**93.26**	**15.76**	**2.47**	**0.98**	-
$sd(\hat{\theta})$		*0.27*	*4.98*	*4.26*	*0.80*	*0.13*	*1.44*	*0.18*	*11.33*	*2.16*	*0.66*	*0.11*	-
$\hat{\boldsymbol{\theta}}$	**0.871** [0.5]	**1.73**	**58.86**	**21.52**	**2.74**	**1.28**	**4.91**	**1.42**	**95.48**	**16.22**	**2.43**	**0.99**	-
$sd(\hat{\theta})$		*0.27*	*5.18*	*4.35*	*0.79*	*0.13*	*1.43*	*0.19*	*11.84*	*2.22*	*0.64*	*0.11*	-
$\hat{\boldsymbol{\theta}}$	**0.890** [0.6]	**1.73**	**59.80**	**21.99**	**2.73**	**1.28**	**4.89**	**1.43**	**97.46**	**16.64**	**2.40**	**0.99**	-
$sd(\hat{\theta})$		*0.27*	*5.37*	*4.44*	*0.78*	*0.13*	*1.42*	*0.19*	*12.31*	*2.27*	*0.63*	*0.10*	-
$\hat{\boldsymbol{\theta}}$	**0.909** [0.7]	**1.73**	**60.69**	**22.45**	**2.72**	**1.28**	**4.88**	**1.44**	**99.31**	**17.05**	**2.37**	**0.99**	-
$sd(\hat{\theta})$		*0.27*	*5.56*	*4.53*	*0.78*	*0.13*	*1.41*	*0.19*	*12.75*	*2.33*	*0.61*	*0.10*	-
$\hat{\boldsymbol{\theta}}$	**0.928** [0.8]	**1.74**	**61.60**	**22.92**	**2.70**	**1.28**	**4.86**	**1.44**	**101.16**	**17.48**	**2.34**	**1.00**	-
$sd(\hat{\theta})$		*0.27*	*5.75*	*4.62*	*0.77*	*0.13*	*1.40*	*0.19*	*13.21*	*2.38*	*0.60*	*0.10*	-
$\hat{\boldsymbol{\theta}}$	**0.950** [0.9]	**1.74**	**62.62**	**23.45**	**2.69**	**1.28**	**4.84**	**1.45**	**103.21**	**17.97**	**2.30**	**1.00**	-
$sd(\hat{\theta})$		*0.27*	*5.97*	*4.72*	*0.76*	*0.13*	*1.39*	*0.19*	*13.73*	*2.44*	*0.58*	*0.10*	-
$\hat{\boldsymbol{\theta}}$	**0.983** [0.99]	**1.74**	**64.12**	**24.24**	**2.68**	**1.28**	**4.81**	**1.47**	**106.16**	**18.70**	**2.26**	**1.00**	-
$sd(\hat{\theta})$		*0.27*	*6.30*	*4.86*	*0.75*	*0.13*	*1.38*	*0.19*	*14.50*	*2.54*	*0.56*	*0.10*	-

C4 Second-degree stochastic dominance analysis

Table C.2: Mean reliability $\mu_{reliab.}$, risk adjusted mean $\mu'_{reliab.}$, and corresponding ranks (2010-2050).

	Alternative	$A_{f2\%}$	$A_{f1.5\%}$	$A_{a2\%}$	$A_{a1.5\%}$	A_{cyc80}	A_{f2+}	$A_{f1\%}$	$A_{a1\%}$	$A_{f0.5\%}$	$A_{a0.5\%}$	A_{f3+}	$A_{fr2\%}$	$A_{fr1.5\%}$	$A_{fr1\%}$	A_{f4+}	A_{f5+}	A_{cyc100}	A_{ref}
Boom	$\mu_{reliab.}$	0.9967	0.9961	0.996	0.9956	0.9954	0.9954	0.9953	0.9951	0.9944	0.9943	0.9941	0.9936	0.9936	0.9936	0.9935	0.9933	0.9932	0.9931
	rank($\mu_{reliab.}$)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	$\mu'_{reliab.}$	0.9985	0.9982	0.998	0.9978	0.9976	0.9975	0.9976	0.9975	0.9971	0.997	0.9968	0.9966	0.9966	0.9966	0.9965	0.9964	0.9964	0.9963
	rank($\mu'_{reliab.}$)	1	2	3	4	6	8	5	7	9	10	11	12	13	14	15	17	16	18
Doom	$\mu_{reliab.}$	0.9954	0.9945	0.9942	0.9931	0.9917	0.9918	0.993	0.9915	0.9902	0.99	0.9894	0.9875	0.9875	0.9875	0.9879	0.9871	0.9887	0.9864
	rank($\mu_{reliab.}$)	1	2	3	4	7	6	5	8	9	10	11	14	15	16	13	17	12	18
	$\mu'_{reliab.}$	0.9967	0.9958	0.9956	0.9943	0.9928	0.9926	0.9944	0.9926	0.9913	0.9912	0.9904	0.9891	0.9891	0.989	0.9892	0.9886	0.9902	0.9882
	rank($\mu'_{reliab.}$)	1	2	3	5	6	8	4	7	9	10	11	14	15	16	13	17	12	18
Qual. of life	$\mu_{reliab.}$	0.9965	0.996	0.9959	0.9954	0.9946	0.9947	0.9952	0.9947	0.9939	0.9937	0.9933	0.9927	0.9927	0.9926	0.9925	0.9922	0.9926	0.9919
	rank($\mu_{reliab.}$)	1	2	3	4	8	7	5	6	9	10	11	12	13	15	16	17	14	18
	$\mu'_{reliab.}$	0.9981	0.9977	0.9976	0.9971	0.996	0.9959	0.997	0.9964	0.9957	0.9954	0.9948	0.9942	0.9942	0.9942	0.9942	0.9938	0.9946	0.9936
	rank($\mu'_{reliab.}$)	1	2	3	4	7	8	5	6	9	10	11	13	14	15	16	17	12	18
Status quo	$\mu_{reliab.}$	0.9954	0.9945	0.9942	0.9931	0.9917	0.9918	0.993	0.9915	0.9902	0.99	0.9894	0.9875	0.9875	0.9875	0.9879	0.9871	0.9887	0.9864
	rank($\mu_{reliab.}$)	1	2	3	4	7	6	5	8	9	10	11	14	15	16	13	17	12	18
	$\mu'_{reliab.}$	0.9967	0.9958	0.9956	0.9943	0.9928	0.9926	0.9944	0.9926	0.9913	0.9912	0.9904	0.9891	0.9891	0.989	0.9892	0.9886	0.9902	0.9882
	rank($\mu'_{reliab.}$)	1	2	3	5	6	8	4	7	9	10	11	14	15	16	13	17	12	18

Table C.3: Mean intergenerational equity (rehabilitation) µrehab., risk adjusted mean µ'rehab., and corresponding ranks (2010-2050).

	Alternative	$A_{f2\%}$	$A_{f1.5\%}$	$A_{a2\%}$	$A_{a1.5\%}$	A_{cyc80}	A_{f2+}	$A_{f1\%}$	$A_{a1\%}$	$A_{f0.5\%}$	$A_{a0.5\%}$	A_{f3+}	$A_{fr2\%}$	$A_{fr1.5\%}$	$A_{fr1\%}$	A_{f4+}	A_{f5+}	A_{cyc100}	A_{ref}
Boom	$\mu_{rehab.}$	0.5217	0.4663	0.4334	0.3813	0.2585	0.2901	0.3791	0.3152	0.2533	0.2109	0.1367	0.0263	0.0246	0.0221	0.0588	0.0226	0.0659	0.0000
	rank($\mu_{rehab.}$)	1	2	3	4	8	7	5	6	9	10	11	14	15	17	13	16	12	18
	$\mu'_{rehab.}$	0.6453	0.5860	0.5373	0.4767	0.3291	0.3587	0.4848	0.4013	0.3365	0.2802	0.1755	0.0363	0.0336	0.0299	0.0783	0.0316	0.1118	0.0000
	rank($\mu'_{rehab.}$)	1	2	3	5	9	7	4	6	8	10	11	14	15	17	13	16	12	18
Doom	$\mu_{rehab.}$	0.6388	0.5722	0.5310	0.4415	0.3460	0.3862	0.4626	0.3293	0.2553	0.2122	0.1962	0.0208	0.0204	0.0196	0.0905	0.0375	0.1295	0.0000
	rank($\mu_{rehab.}$)	1	2	3	5	7	6	4	8	9	10	11	15	16	17	13	14	12	18
	$\mu'_{rehab.}$	0.7610	0.6997	0.6585	0.5503	0.4051	0.4502	0.5911	0.4058	0.3333	0.2756	0.2532	0.0238	0.0234	0.0228	0.1256	0.0555	0.2091	0.0000
	rank($\mu'_{rehab.}$)	1	2	3	5	8	6	4	7	9	10	11	15	16	17	13	14	12	18
Qual. of life	$\mu_{rehab.}$	0.6356	0.5756	0.5508	0.4776	0.3251	0.3637	0.4803	0.3731	0.2984	0.2289	0.1825	0.0286	0.0277	0.0259	0.0834	0.0343	0.1157	0.0000
	rank($\mu_{rehab.}$)	1	2	3	5	8	7	4	6	9	10	11	15	16	17	13	14	12	18
	$\mu'_{rehab.}$	0.7604	0.7066	0.6856	0.6057	0.3812	0.4224	0.6131	0.4796	0.4009	0.2999	0.2333	0.0376	0.0363	0.0340	0.1146	0.0503	0.1880	0.0000
	rank($\mu'_{rehab.}$)	1	2	3	5	9	7	4	6	8	10	11	15	16	17	13	14	12	18
Status quo	$\mu_{rehab.}$	0.6388	0.5722	0.5310	0.4415	0.3460	0.3862	0.4626	0.3293	0.2553	0.2122	0.1962	0.0208	0.0204	0.0196	0.0905	0.0375	0.1295	0.0000
	rank($\mu_{rehab.}$)	1	2	3	5	7	6	4	8	9	10	11	15	16	17	13	14	12	18
	$\mu'_{rehab.}$	0.7610	0.6997	0.6585	0.5503	0.4051	0.4502	0.5911	0.4058	0.3333	0.2756	0.2532	0.0238	0.0234	0.0228	0.1256	0.0555	0.2091	0.0000
	rank($\mu'_{rehab.}$)	1	2	3	5	8	6	4	7	9	10	11	15	16	17	13	14	12	18

Table C.4: Mean costs (% of average income) $\mu_{cost.}$, risk adjusted mean $\mu'_{cost.}$, and corresponding ranks (2010-2050).

	Alternative	$A_{f2\%}$	$A_{f1.5\%}$	$A_{a2\%}$	$A_{a1.5\%}$	A_{cyc80}	A_{f2+}	$A_{f1\%}$	$A_{a1\%}$	$A_{f0.5\%}$	$A_{a0.5\%}$	A_{f3+}	$A_{fr2\%}$	$A_{fr1.5\%}$	$A_{fr1\%}$	A_{f4+}	A_{f5+}	A_{cyc100}	A_{ref}
Boom	μ_{cost}	0.1697	0.1285	0.1699	0.1279	0.0294	0.0289	0.0859	0.0860	0.0440	0.0441	0.0125	0.1706	0.1277	0.0864	0.0064	0.0038	0.0061	0.0023
	rank(μ_{cost})	16	15	17	14	7	6	10	11	8	9	5	18	13	12	4	2	3	1
	μ'_{cost}	0.0973	0.0744	0.0976	0.0737	0.0051	0.0065	0.0497	0.0498	0.0260	0.0261	0.0048	0.0984	0.0735	0.0504	0.0032	0.0023	0.0019	0.0014
	rank(μ'_{cost})	16	15	17	14	6	7	10	11	8	9	5	18	13	12	4	3	2	1
Doom	μ_{cost}	0.2824	0.0021	0.0028	0.0022	0.0015	0.0008	0.0015	0.0015	0.0008	0.0008	0.0006	0.0029	0.0022	0.0015	0.0004	0.0003	0.0007	0.0001
	rank(μ_{cost})	16	13	17	14	10	8	9	11	6	7	4	18	15	12	3	2	5	1
	μ'_{cost}	0.2519	0.1908	0.2532	0.1923	0.0573	0.0576	0.1303	0.1315	0.0705	0.0709	0.0449	0.2582	0.1963	0.1345	0.0280	0.0184	0.0147	0.0107
	rank(μ'_{cost})	16	13	17	14	6	7	10	11	8	9	5	18	15	12	4	3	2	1
Qual. of life	μ_{cost}	0.2064	0.1558	0.2069	0.1564	0.0689	0.0482	0.1054	0.1061	0.0555	0.0559	0.0286	0.2101	0.1591	0.1081	0.0177	0.0116	0.0289	0.0064
	rank(μ_{cost})	16	13	17	14	9	6	10	11	7	8	4	18	15	12	3	2	5	1
	μ'_{cost}	0.1909	0.1444	0.1916	0.1451	0.0225	0.0263	0.0981	0.0986	0.0521	0.0524	0.0213	0.1944	0.1474	0.1004	0.0137	0.0094	0.0068	0.0055
	rank(μ'_{cost})	16	13	17	14	6	7	10	11	8	9	5	18	15	12	4	3	2	1
Status quo	μ_{cost}	0.1812	0.1371	0.1821	0.1383	0.0933	0.0581	0.0935	0.0947	0.0507	0.0511	0.0377	0.1865	0.1419	0.0974	0.0243	0.0162	0.0426	0.0084
	rank(μ_{cost})	16	13	17	14	9	8	10	11	6	7	4	18	15	12	3	2	5	1
	μ'_{cost}	0.1755	0.1326	0.1763	0.1339	0.0336	0.0374	0.0902	0.0918	0.0490	0.0495	0.0300	0.1818	0.1384	0.0950	0.0186	0.0125	0.0093	0.0074
	rank(μ'_{cost})	16	13	17	14	6	7	10	11	8	9	5	18	15	12	4	3	2	1

C5 MCDA results for all alternatives

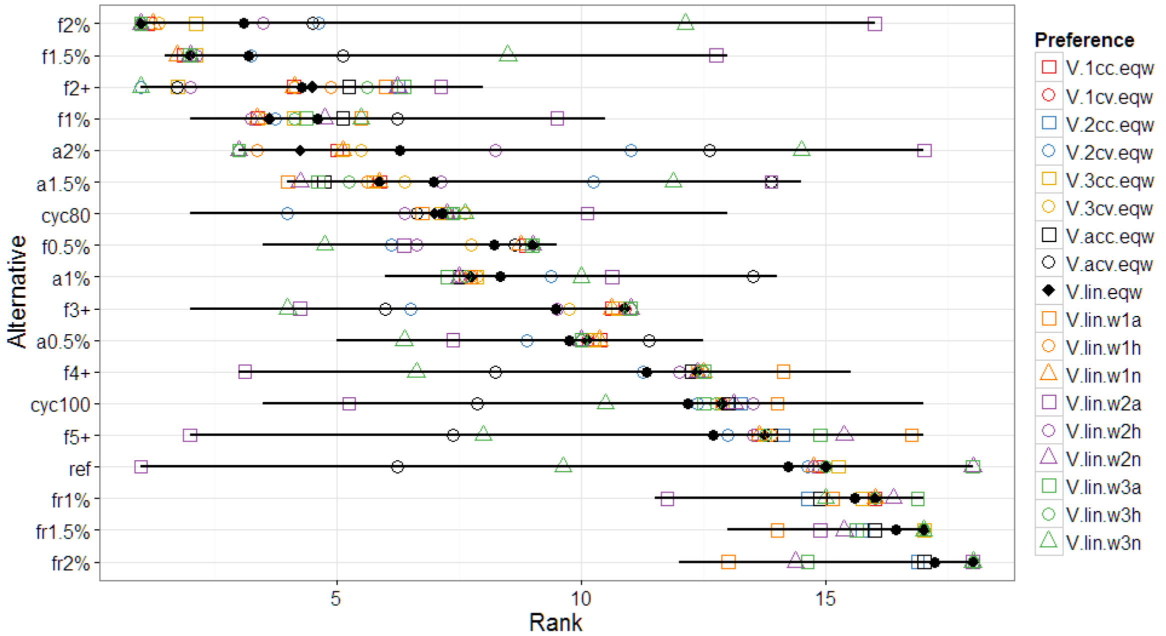


Figure C.1: Sensitivity of the ranking to different weights and value function forms without assumption of any specific risk attitude. The black point and line represent mean rank and rank ranges (minimum and maximum rank) of the outcomes of alternatives. Ranks are aggregated over the four scenarios

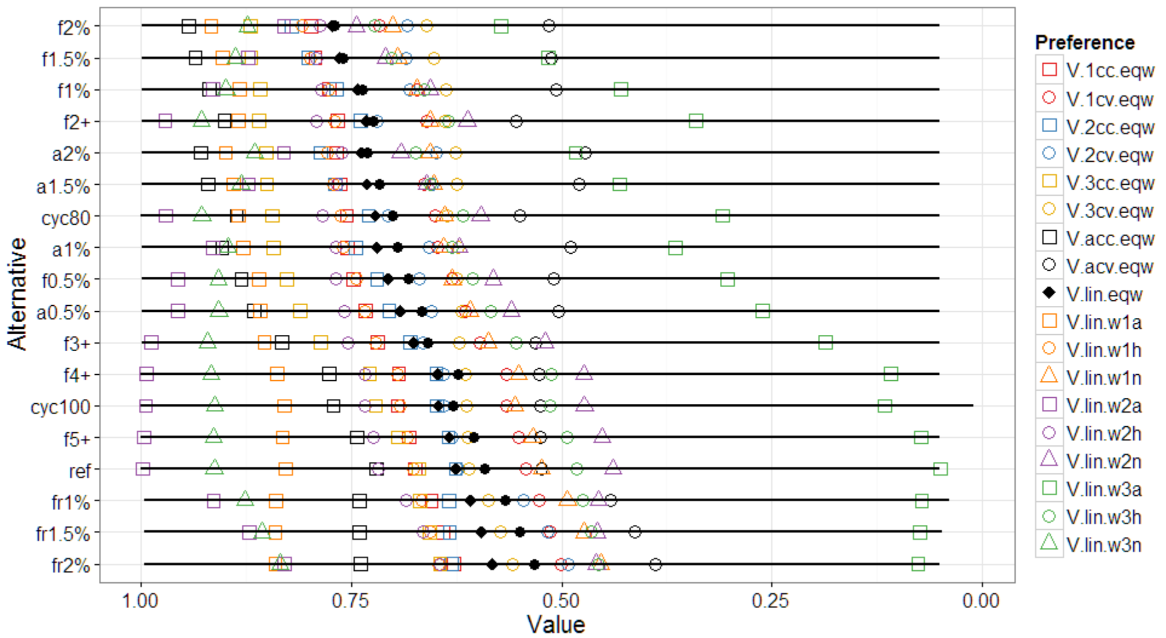


Figure C.2: Sensitivity of the overall value of the alternatives to weight and value function changes without assumption of any specific risk attitude. The black point and line represent mean values and value ranges (absolute minimum and maximum value) of the outcomes of alternatives. Values are aggregated over the four scenarios

C6 Additional figures

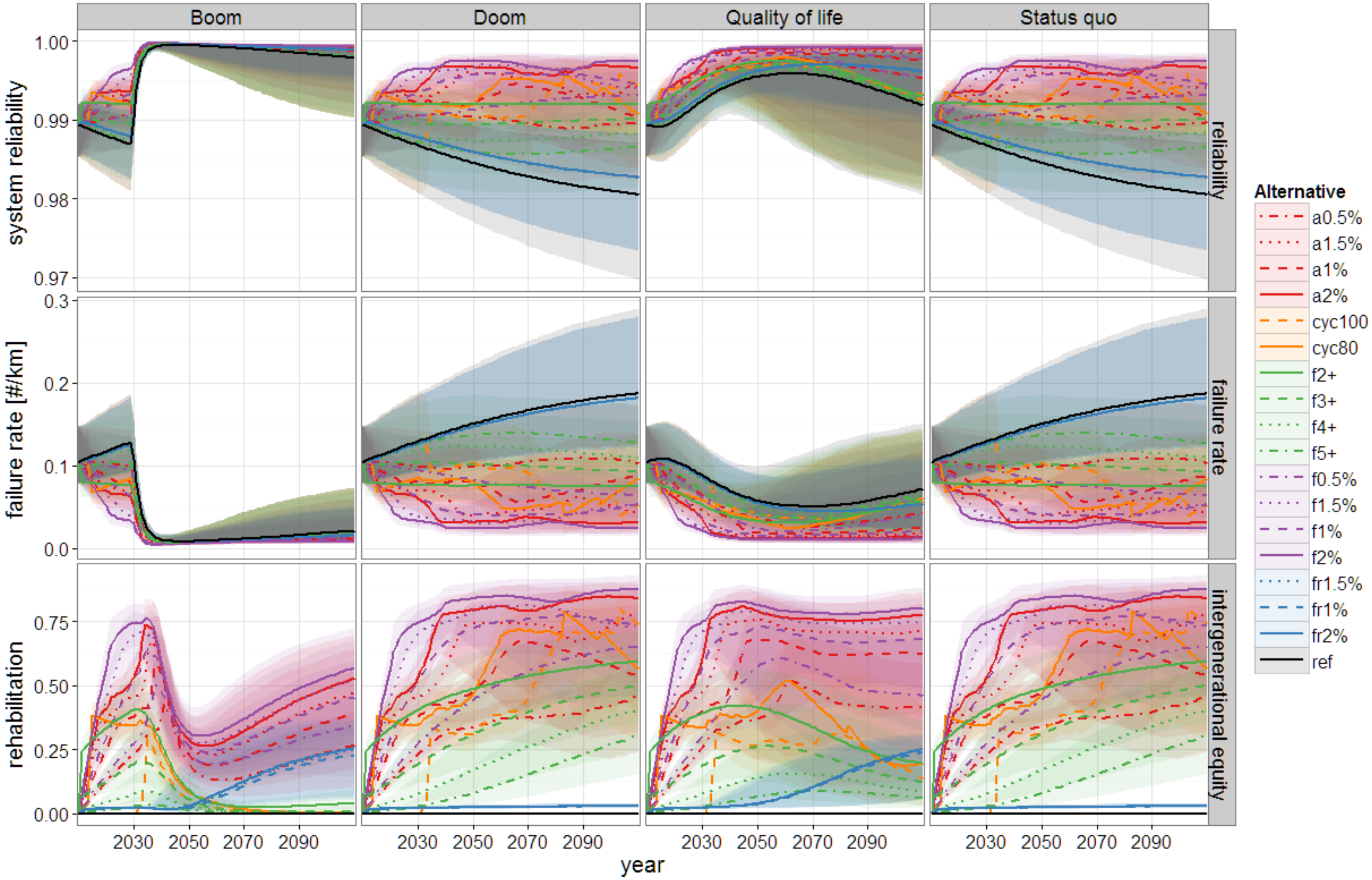


Figure C.3: Outcomes of the alternatives for reliability and intergenerational equity plotted against the development of the failure rate. The strong relationship especially between reliability and failure rate is apparent.

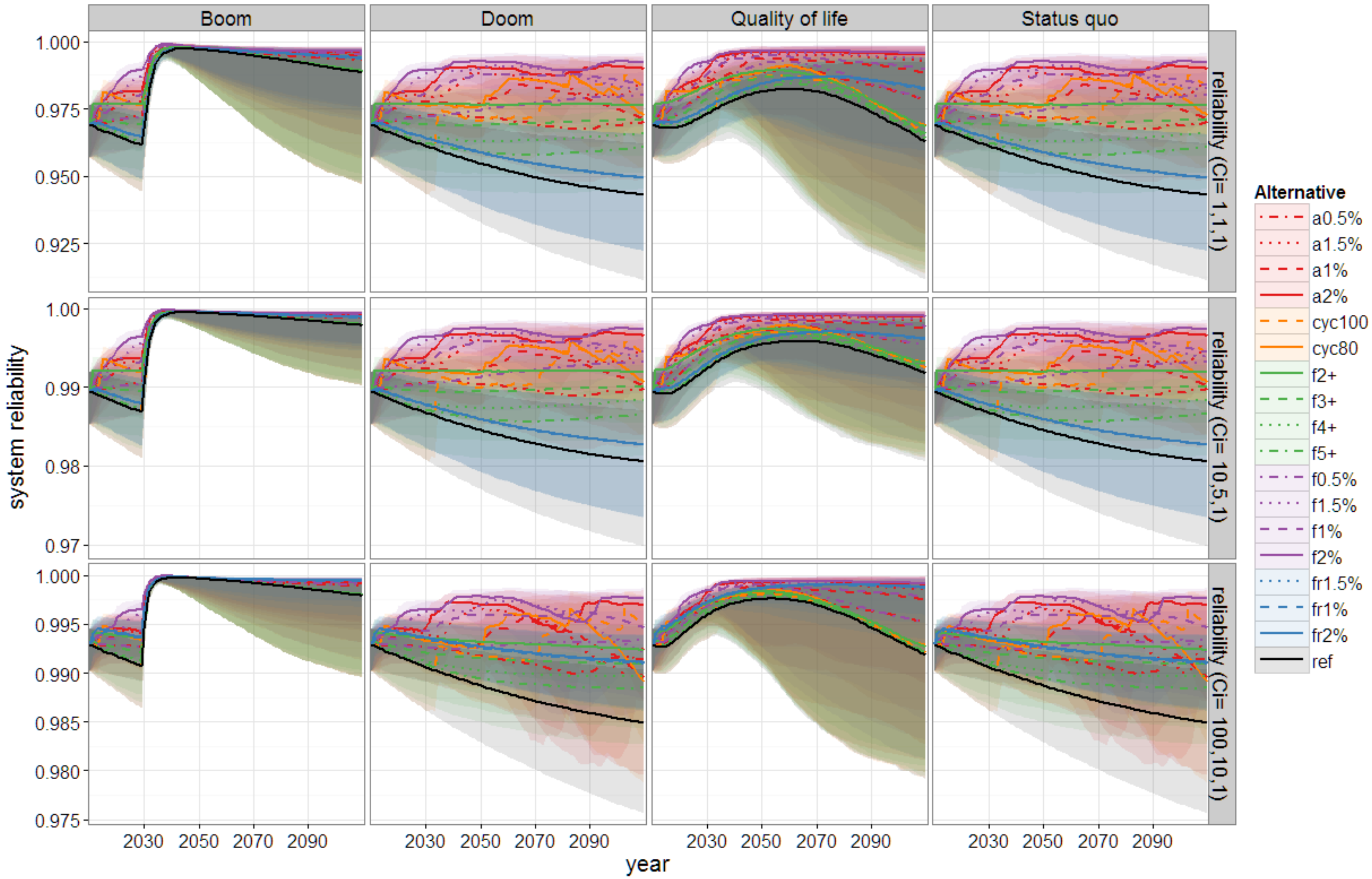


Figure C.4: Reliability under different assumptions for the criticality indices (in following order: ($C_{\geq 250\,mm}$, $C_{150\text{-}250mm}$, $C_{\leq 150mm}$). Note the considerable improvement of $A_{fr2\ \ 1\%}$ (blue lines) with increasing criticality of larger pipes.

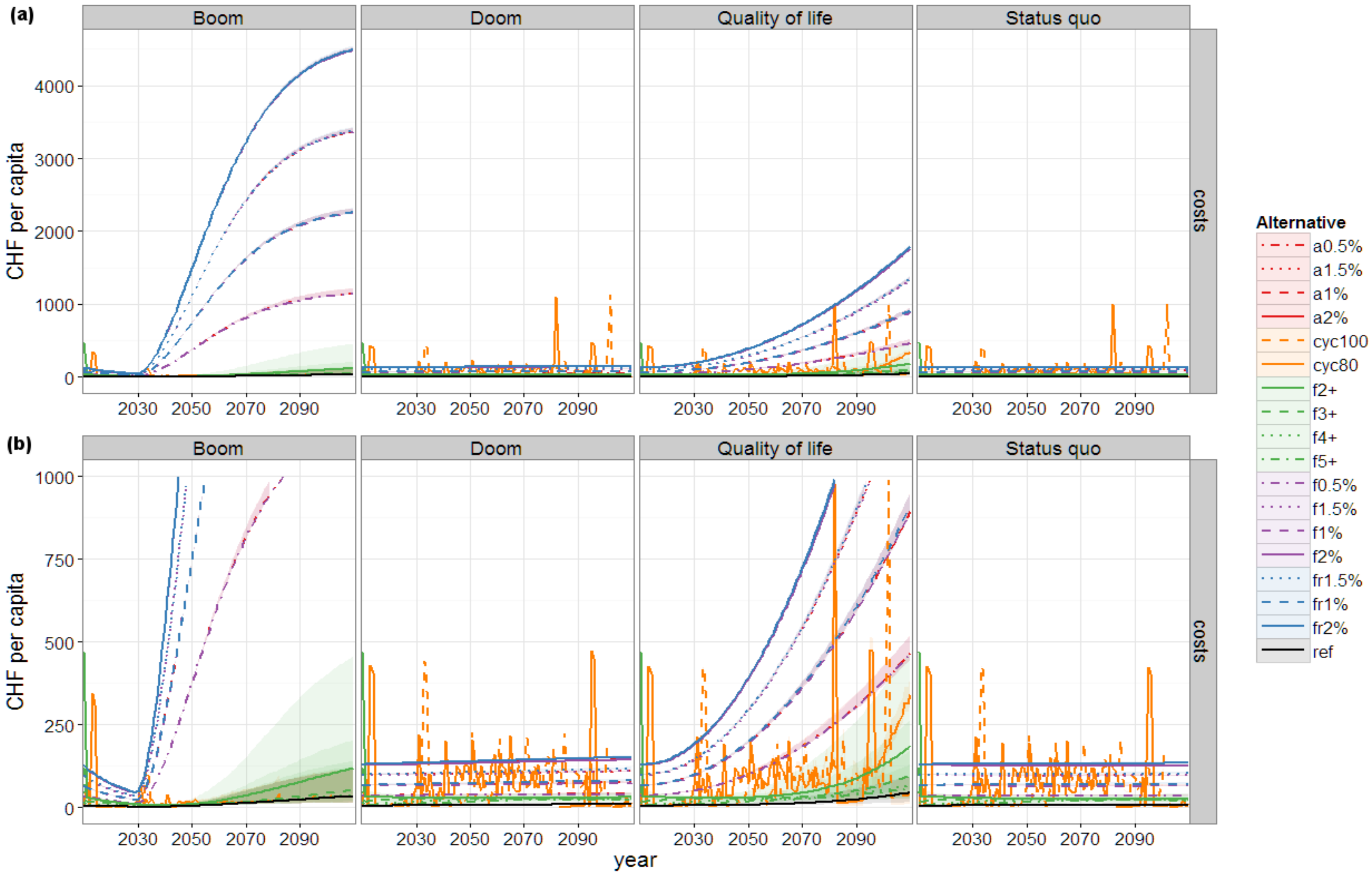


Figure C.26: Development of absolute per capita costs. (a) shows the results on original scale, (b) on a rescaled scale to better demonstrate results <250 CHF per capita. The results are displayed without considering neither discount rates for repair and replacement costs, nor inflation of incomes. Note the strong increase of costs in scenarios with high infrastructure expansion. Note the strong cost increase of alternatives with network-length dependent replacement strategies in scenarios with high infrastructure expansion.

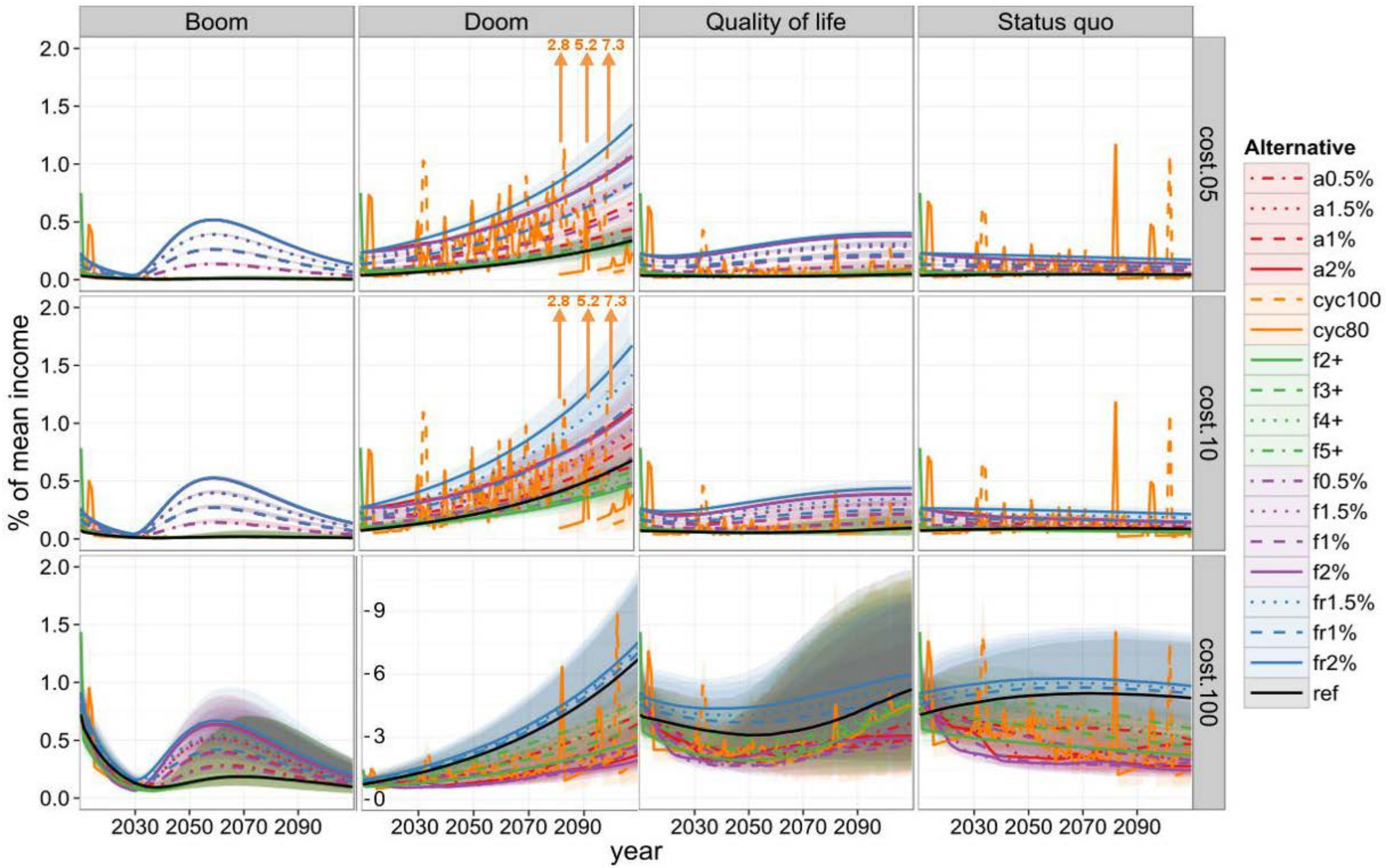


Figure C.6: Sensitivity of the attribute outcomes for costs to different unit cost assumptions. Note the adapted zoom for "cost.100" under the Doom scenario. "Cost.05" stands for five times higher repair costs, "cost.10", and "cost.100" for ten and a hundred times higher repair costs (i.e. 32'500, 65'000, 650'000 CHF per repair respectively). This is equivalent to repair to replacement cost ratios of approx. 1:3, 1:1.5, 1:0.15 while the ratio underlying the assumptions and results presented in the main text is 1:15. Note the increase of uncertainty with increasing repair costs (as only parametric uncertainty of the failure model is propagated).

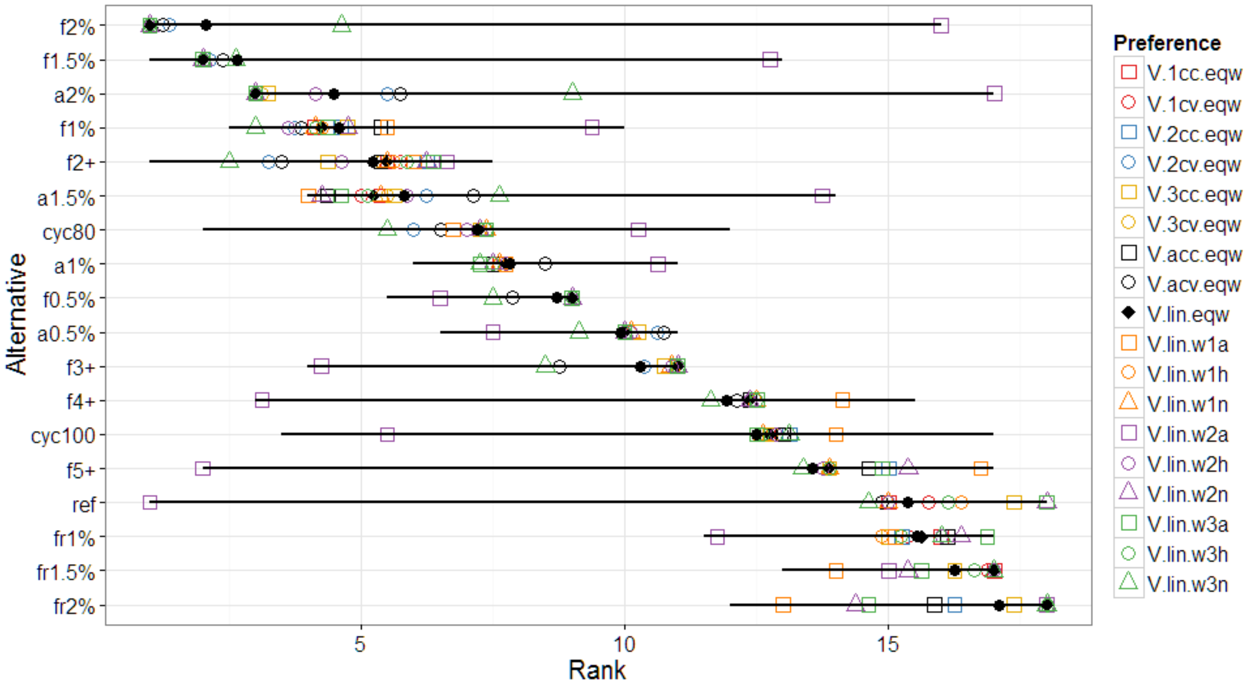


Figure C.7: Sensitivity of the ranking <u>assuming five times higher repair costs</u> to different weights and value function forms without assumption of any specific risk attitude. The black point and line represent mean rank and rank ranges (minimum and maximum rank) of the outcomes of alternatives. Ranks are aggregated over the four scenarios

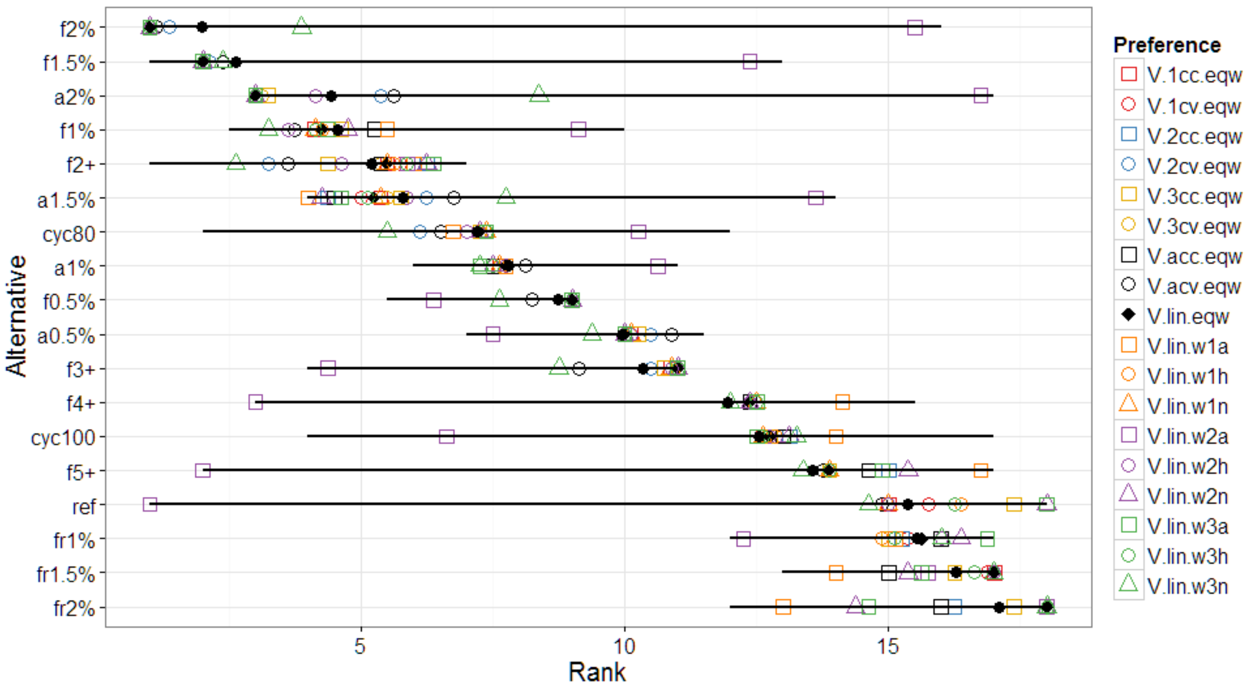


Figure C.8: Sensitivity of the ranking <u>assuming ten times higher repair costs</u> to different weights and value function forms without assumption of any specific risk attitude. The black point and line represent mean rank and rank ranges (minimum and maximum rank) of the outcomes of alternatives. Ranks are aggregated over the four scenarios

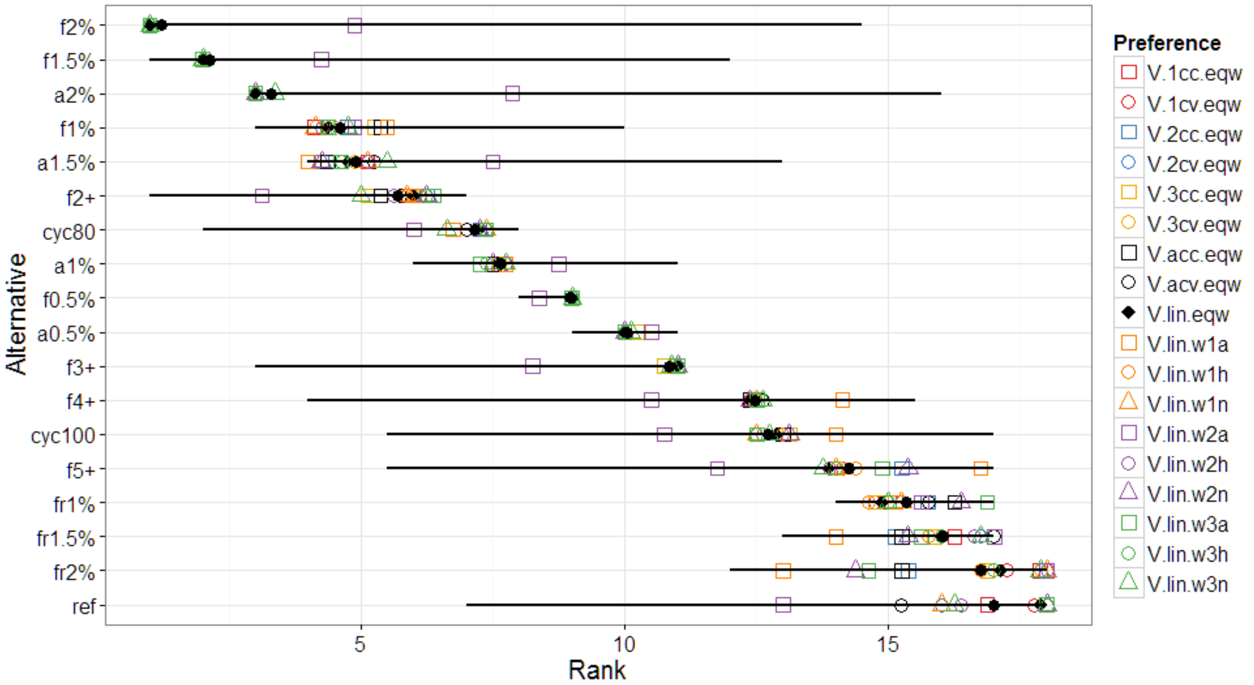


Figure C.9: Sensitivity of the ranking <u>assuming a hundred times higher repair costs</u> to different weights and value function forms without assumption of any specific risk attitude. The black point and line represent mean rank and rank ranges (minimum and maximum rank) of the outcomes of alternatives. Ranks are aggregated over the four scenarios

SI-D) Supporting Information to: 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, $z = 1…90$

D1 Stakeholder identification

- More details about the underlying stakeholder (SH) and social network analysis (SNA) are given in (Lienert et al., submitted). 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 D1.1: 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.

D2 Decision attributes

Description of attributes and quantification

Table D2.1: Overview of the ranges, description, and assessment of the attributes. The distributions and corresponding parameters used are shown in Table D3.2. 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.

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
gwhh	% Utilization of groundwater recharge [%]	+0-180	Raw water can be abstracted from springs and groundwater wells in the 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).	Calculated as groundwater abstraction/groundwater recharge. Groundwater recharge was estimated using the Hydrus1D model for simplified soil profiles, representing 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

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
				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 (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,	Changes in total cell	+0-2	Cell counts are indicators for the amount	Expert assessment. Expert and assessment as for aes_dw, aes_hw, but with an

Short name	**Attribute [unit]**	**Range** (+: best level)	**Description**	**Assessment**
cells_hw	count as indicator of bacterial regrowth [log units]		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 in the microbial community and hence regrowth, which is usually of higher interest than absolute cell counts.	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.
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	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
			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 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	+0-10	This attribute estimates the time each	Calculated. Only applies to decentralized installations in private households which

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
	investment for operation and maintenance by user [h/(inh.*a)]		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.	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 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 ground.	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 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. D3.1 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.

Short name	Attribute [unit]	Range (+: best level)	Description	Assessment
				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.

D3 Decision alternatives

Overview of decision alternatives

Table D3.1: 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

No.	Name	Organization, cooperation, management	Rehabilitation, operation, and maintenance	Water supply and uses	Water sources	Water treatment technology
						a POE treatment module (GAC+UF) before use.
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.

No.	Name	Organization, cooperation, management	Rehabilitation, operation, and maintenance	Water supply and uses	Water sources	Water treatment technology
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 retention	The communities [(b)] remain responsible for a single, integrated wastewater and drinking water sector that jointly operate the water infrastructures, with some services contracted out to private enterprises.	Rehabilitation is done according to prioritization (1 % annual replacement by condition and criticality). Renovation is trenchless. Their operation, maintenance, and inspection is moderate.	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 imported from the regional water supplier.	Groundwater is disinfected (UV treatment), lake water receives a multi-step treatment as today, 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 D3.2: Predictions of the attributes (Tab. D2.1) by alternative and scenario, stated as probability distributions. Explanation of abbreviations: A1a – A9…alternatives; see Table D3.1 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	N(0.713,	N(0.713,	N(0.713,	N(0.0778,	N(0.4,0.1)	N(0.3649,	N(0.55,	N(0.1797,	N(0.67,	N(0.67,	N(0.67,

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
life	0.1783)	0.1783)	0.1783)	0.0194)		0.0912)	0.1375)	0.0449)	0.1675)	0.1675)	0.1675)
DW: Days per year with water quantity limitations [d/a] (vol_dw)											
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)

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
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)											
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)

	A1a	**A1b**	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8a**	**A8b**	**A9**
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)
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	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)

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
life											
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)											
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(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)
Doom	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)
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)

	A1a	A1b	A2	A3	A4	A5	A6	A7	A8a	A8b	A9
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)											
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 = 10'000)

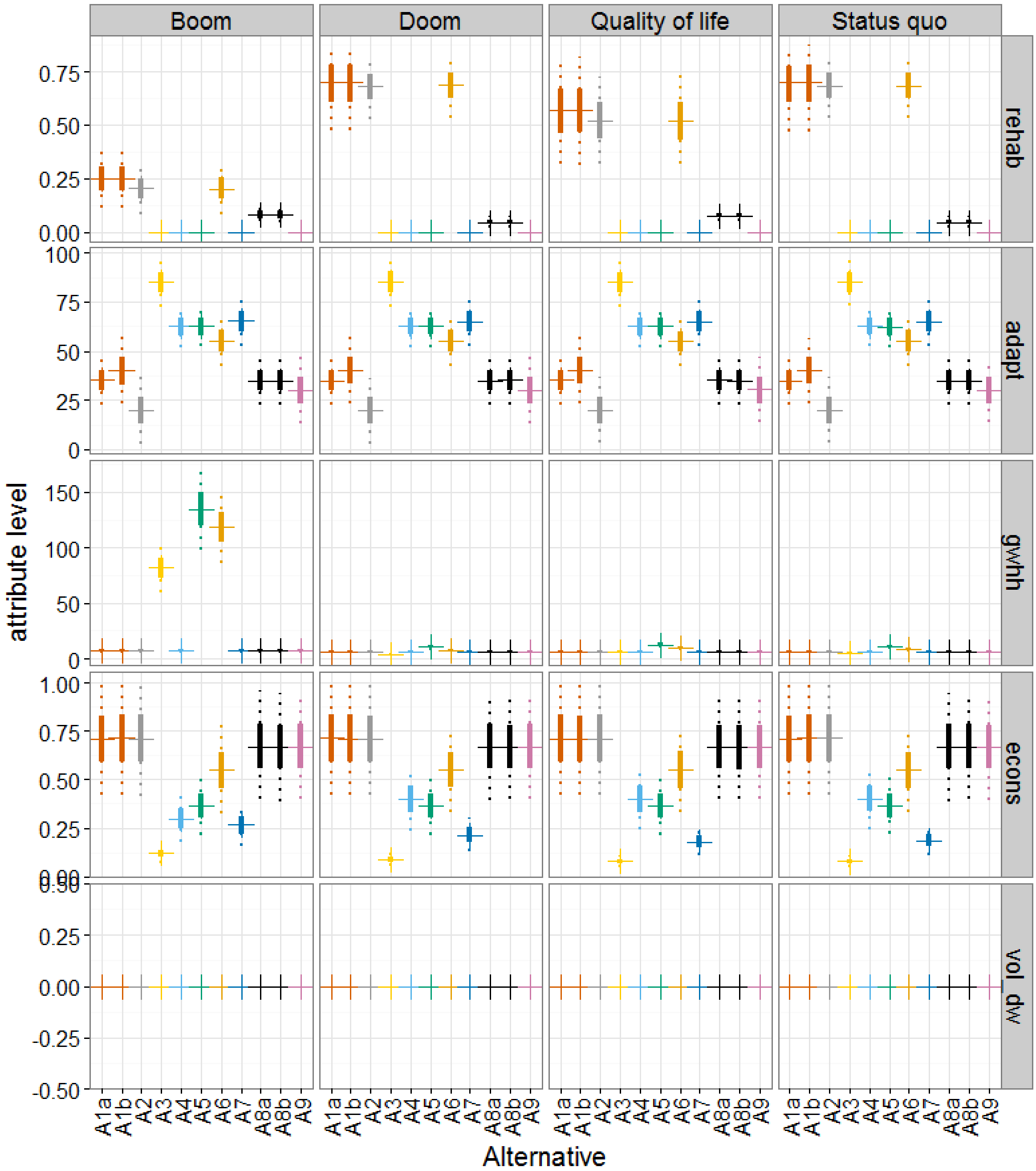


Figure D3.1a: Distribution of attribute levels by alternative (A1a to A9; see Tab. D3.1) under four future scenarios (Boom, Doom, Quality of life, Status quo). Labels on the right correspond to the short names of the attributes as listed in Table D2.1. The predicted attribute levels and their uncertainty are given on the y-axis (see Tab. D3.1 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.

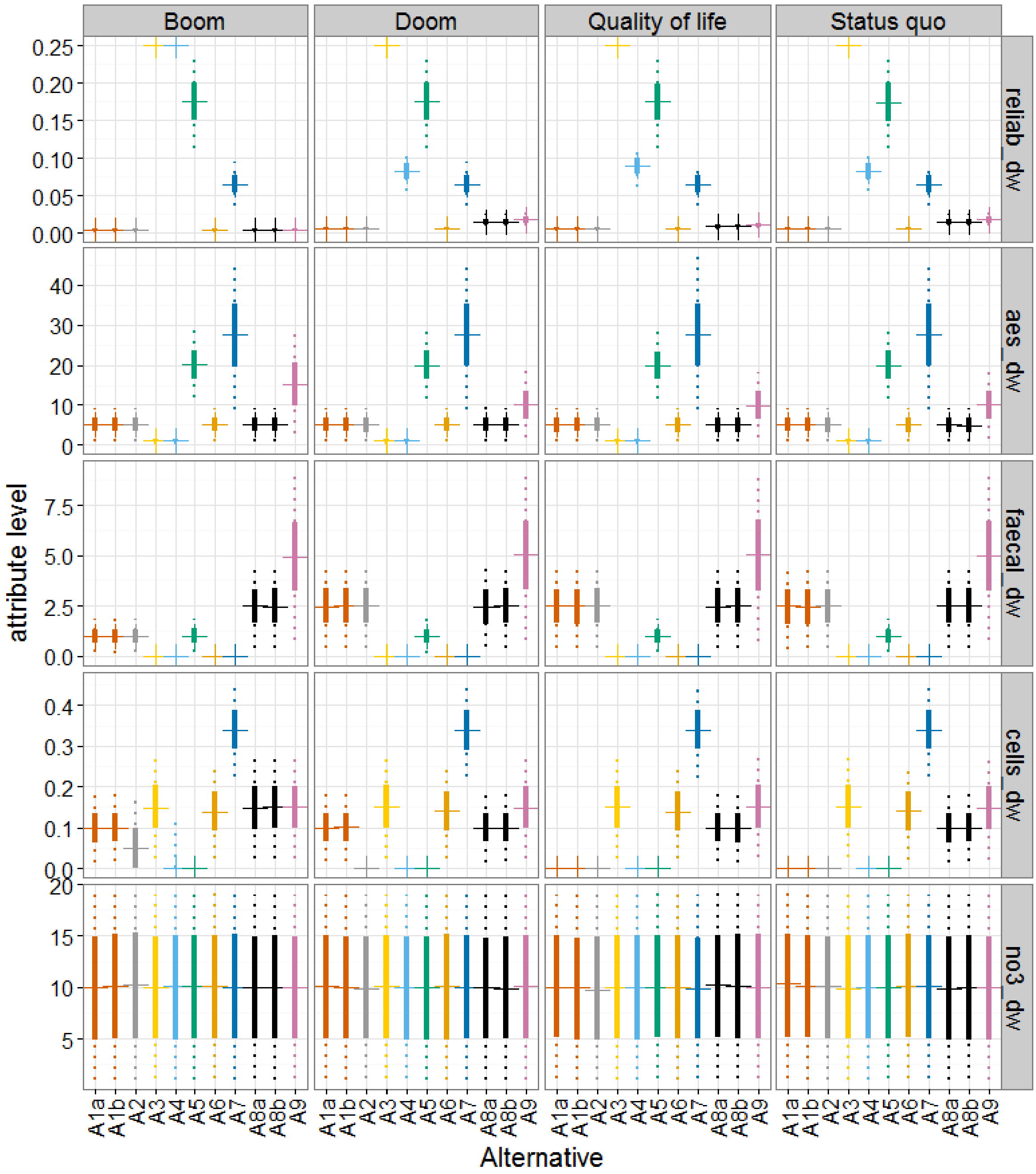


Figure D3.1b For descriptions see Fig. D3.1a.

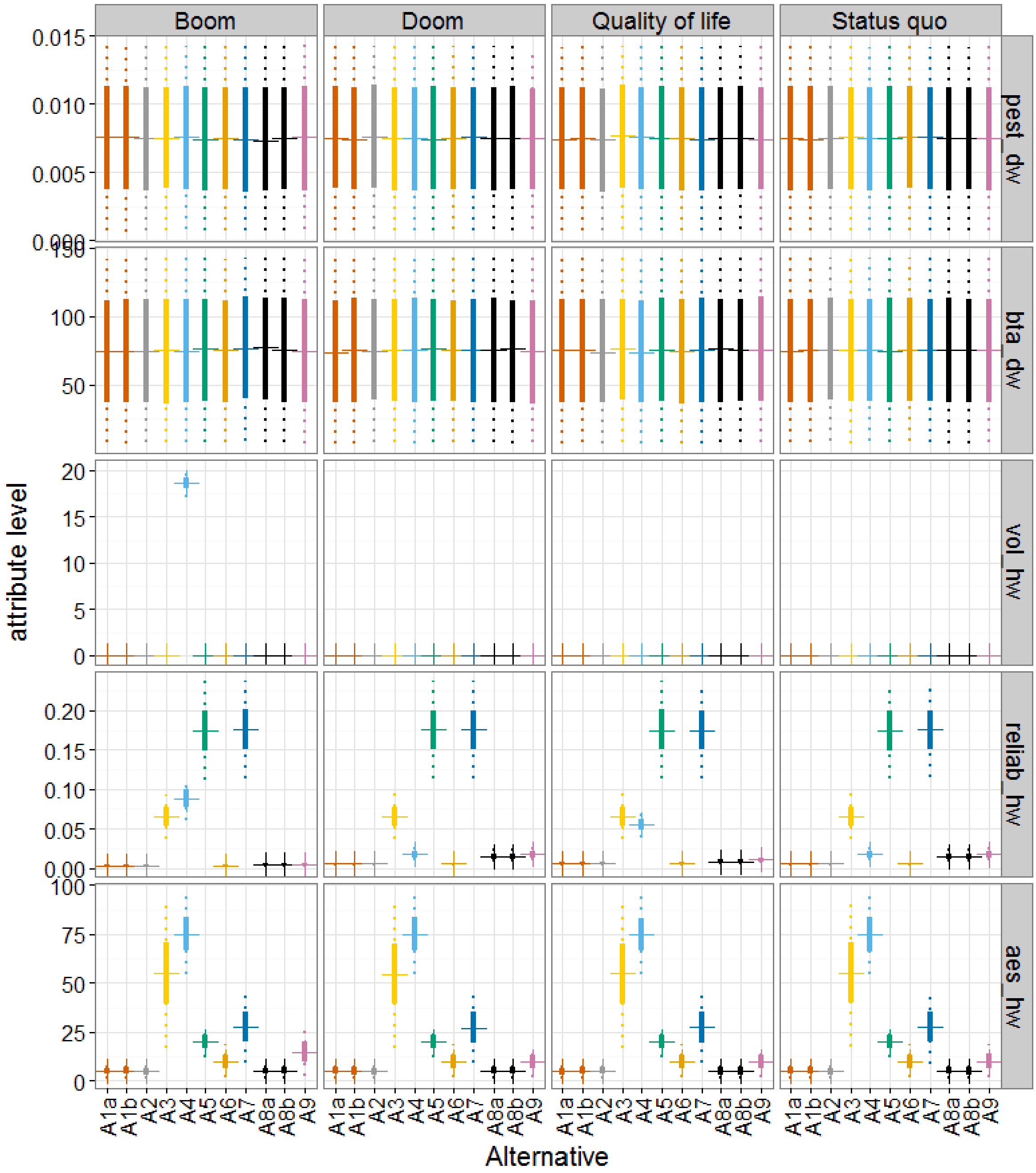


Figure D3.1c For descriptions see Fig. D3.1a.

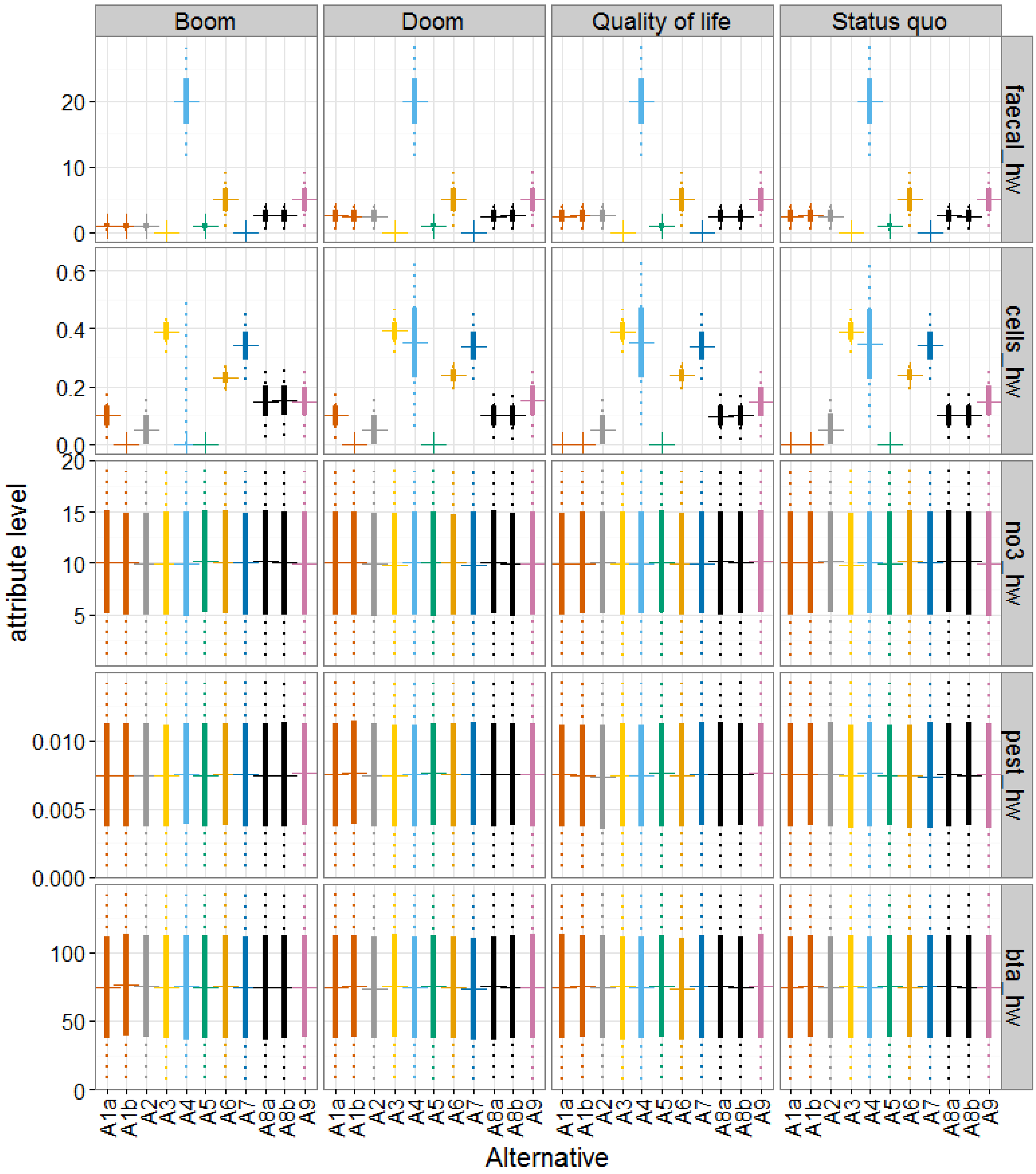


Figure D3.1d For descriptions see Fig. D3.1a.

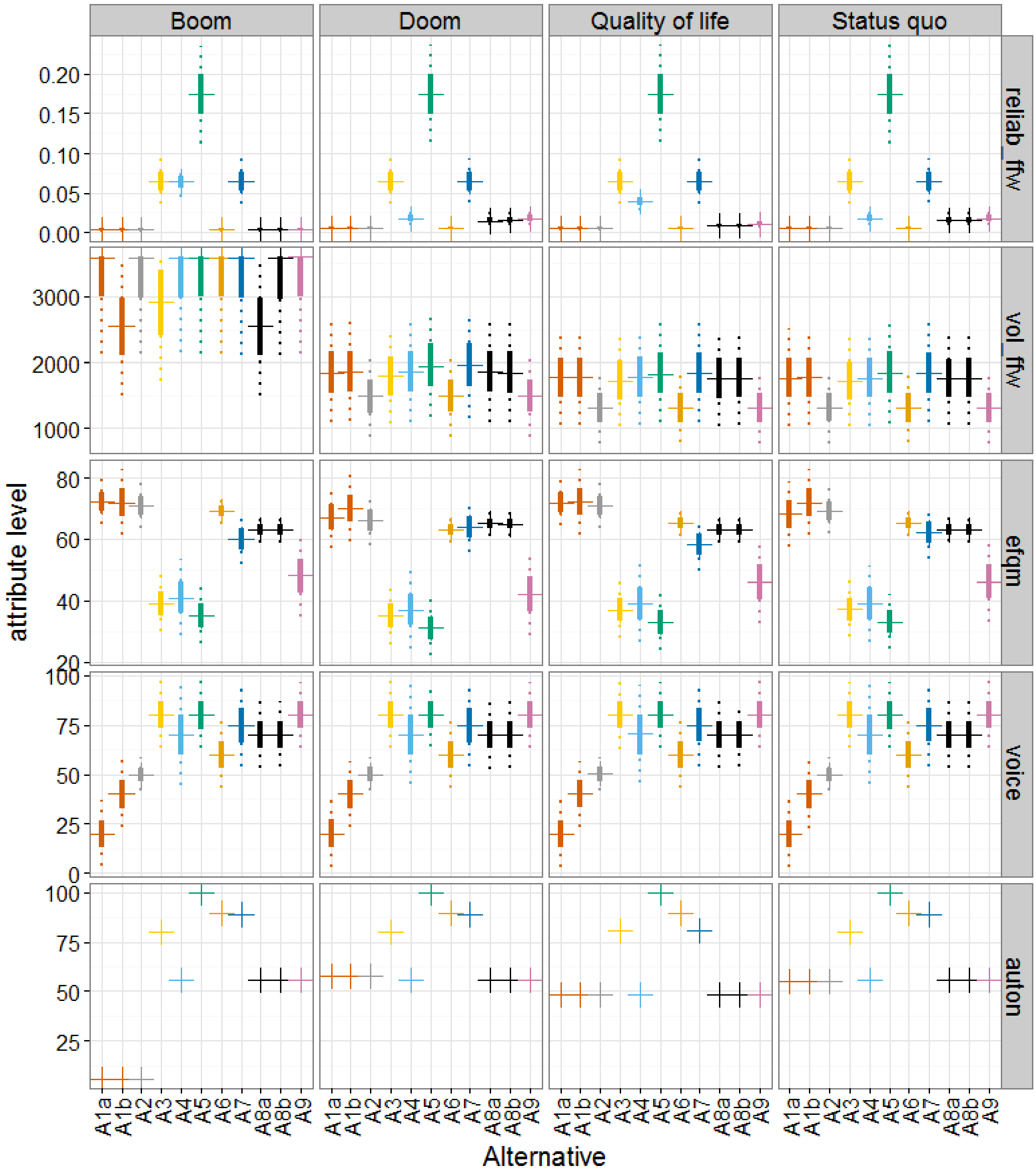


Figure D3.1e For descriptions see Fig. D3.1a.

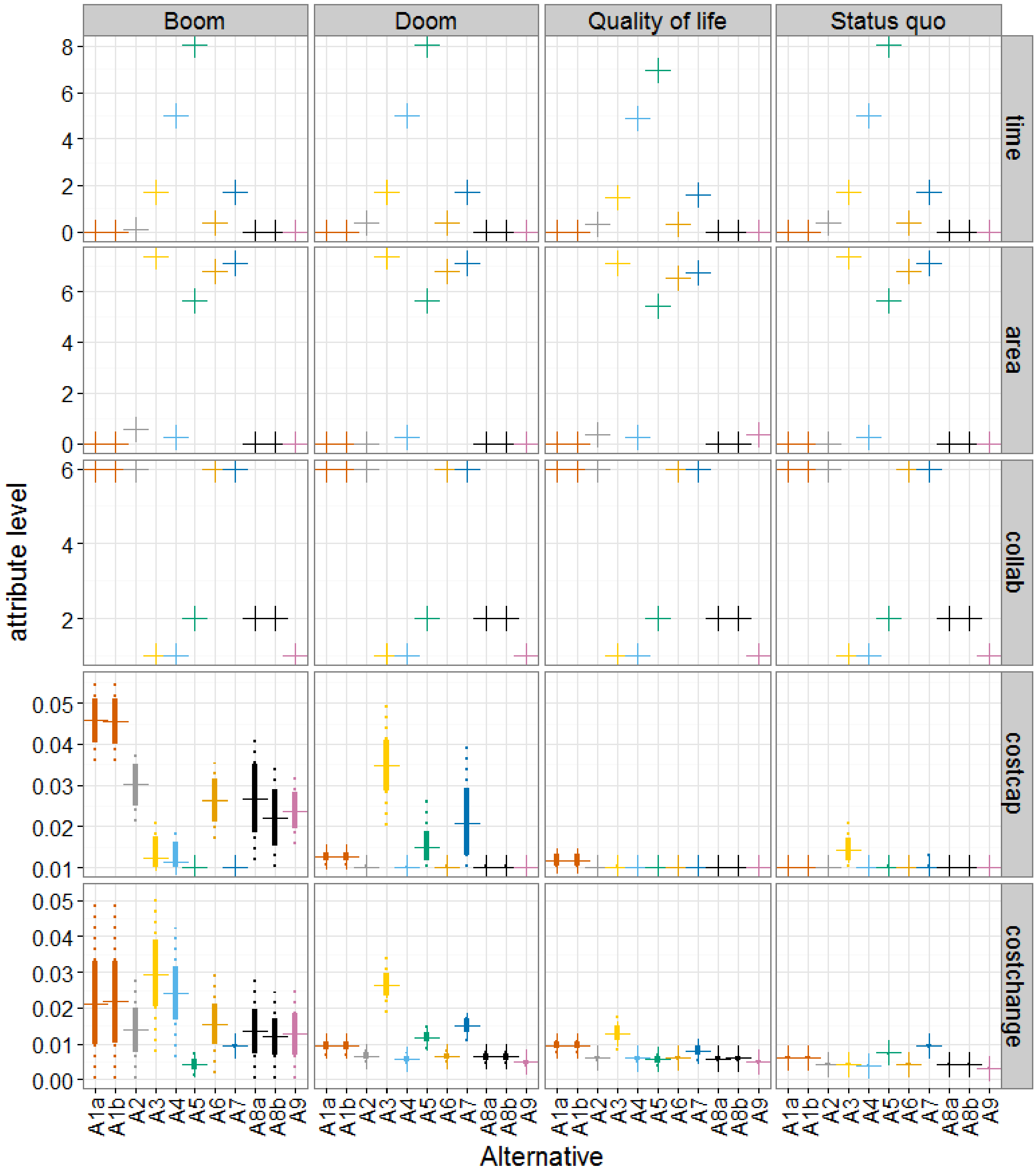


Figure D3.1f For descriptions see Fig. D3.1a.

D4 Stakeholder preferences

Weights

Table D4.1: Elicited weights from face-to-face interviews with ten stakeholders (see Tab. D1.1). 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. D2.1). Objectives receiving zero weight (= irrelevant) are grey shaded. SH = stakeholder, dw = drinking water, hw= household water, ffw = firefighting water.

		SH1	SH2	SH3	SH4	SH5	SH6	SH7	SH8	SH9	SH10	over-all
Objective	**Weight no.**	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 0.17 0.21	0.16 0.19 0.23	0.10 0.12 0.14	0.26 0.29 0.31	0.25 0.30 0.34	0.19 0.24 0.29	0.13 0.17 0.21	0.18 0.20 0.23	0.17 0.20 0.23	0.00 0.00 0.00	0.00 0.19 0.34
Resources & groundwater protection	**w.2**	0.20 0.24 0.29	0.19 0.22 0.26	0.21 0.24 0.28	0.21 0.23 0.26	0.06 0.10 0.14	0.13 0.17 0.21	0.21 0.25 0.29	0.29 0.31 0.34	0.12 0.15 0.18	0.40 0.43 0.48	0.06 0.24 0.48
Water supply	**w.3**	0.31 0.34 0.38	0.26 0.28 0.30	0.31 0.34 0.37	0.33 0.36 0.38	0.29 0.33 0.37	0.30 0.34 0.38	0.25 0.28 0.31	0.34 0.37 0.39	0.23 0.24 0.27	0.35 0.39 0.43	0.23 0.33 0.43
Social acceptance	**.w4**	0.00 0.00 0.00	0.11 0.13 0.15	0.06 0.08 0.11	0.02 0.04 0.06	0.03 0.05 0.07	0.09 0.12 0.15	0.08 0.11 0.15	0.02 0.03 0.04	0.17 0.20 0.23	0.00 0.00 0.00	0.00 0.08 0.23
Costs	**w.5**	0.20 0.24 0.29	0.16 0.18 0.21	0.18 0.22 0.26	0.07 0.09 0.11	0.19 0.23 0.28	0.10 0.14 0.18	0.16 0.19 0.24	0.07 0.09 0.11	0.20 0.22 0.23	0.13 0.17 0.22	0.07 0.18 0.29
Rehabilita-tion burden	**w.1.1**	0.59 0.63 0.67	0.63 0.67 0.71	0.23 0.26 0.29	0.54 0.56 0.57	0.53 0.56 0.59	0.56 0.59 0.63	0.41 0.44 0.47	0.77 0.80 0.83	0.63 0.67 0.71	0.00 0.00 0.00	0.00 0.52 0.83
Flexibility	**w.1.2**	0.33 0.38 0.41	0.29 0.33 0.38	0.71 0.74 0.77	0.43 0.44 0.46	0.41 0.44 0.47	0.38 0.41 0.44	0.53 0.56 0.59	0.17 0.20 0.23	0.29 0.33 0.38	0.00 0.00 0.00	0.00 0.38 0.77
Groundwater protection	**w.2.1**	0.63 0.67 0.71	0.67 0.71 0.77	0.71 0.74 0.77	0.71 0.74 0.77	0.38 0.41 0.44	0.63 0.67 0.71	0.77 0.80 0.83	0.63 0.67 0.71	0.80 0.83 0.87	1.00 1.00 1.00	0.38 0.73 1.00
Energy consumption	**w.2.2**	0.29 0.33 0.38	0.23 0.29 0.33	0.23 0.26 0.29	0.23 0.26 0.29	0.56 0.59 0.63	0.29 0.33 0.38	0.17 0.20 0.23	0.29 0.33 0.38	0.13 0.17 0.20	0.00 0.00 0.00	0.00 0.28 0.63
Dw supply	**w.3.1**	0.36 0.38 0.42	0.40 0.43 0.45	0.48 0.51 0.54	0.48 0.50 0.53	0.28 0.32 0.36	0.67 0.74 0.83	0.33 0.36 0.38	0.59 0.67 0.77	0.42 0.45 0.50	0.36 0.37 0.38	0.28 0.48 0.83
Hw supply	**w.3.2**	0.27 0.31 0.35	0.29 0.32 0.35	0.26 0.29 0.33	0.08 0.10 0.13	0.37 0.40 0.43	0.07 0.11 0.15	0.29 0.32 0.36	0.23 0.33 0.41	0.32 0.36 0.41	0.36 0.37 0.38	0.07 0.29 0.43
Ffw supply	**w.3.3**	0.27 0.31 0.35	0.22 0.26 0.29	0.18 0.20 0.23	0.38 0.40 0.43	0.24 0.28 0.32	0.08 0.15 0.21	0.29 0.32 0.36	0.00 0.00 0.00	0.14 0.18 0.23	0.23 0.26 0.29	0.00 0.24 0.43
Dw quantity	**w.3.1.1**	0.17 0.22 0.26	0.25 0.27 0.29	0.18 0.20 0.23	0.18 0.22 0.26	0.26 0.31 0.36	0.21 0.23 0.25	0.24 0.27 0.31	0.00 0.04 0.08	0.21 0.25 0.29	0.17 0.21 0.25	0.00 0.22 0.36
Dw reliability	**w.3.1.2**	0.40 0.43 0.48	0.29 0.31 0.33	0.26 0.29 0.33	0.28 0.32 0.36	0.29 0.33 0.37	0.33 0.35 0.38	0.31 0.33 0.36	0.15 0.22 0.29	0.29 0.33 0.38	0.38 0.42 0.45	0.15 0.33 0.48
Dw quality	**w.3.1.3**	0.30 0.35 0.39	0.40 0.42 0.43	0.48 0.51 0.54	0.43 0.46 0.50	0.33 0.36 0.40	0.40 0.42 0.43	0.37 0.39 0.42	0.67 0.74 0.83	0.38 0.42 0.45	0.33 0.38 0.42	0.30 0.45 0.83
Hw quantity	**w.3.2.1**	0.27 0.32 0.36	0.27 0.30 0.32	0.18 0.20 0.23	0.23 0.26 0.29	0.38 0.43 0.48	0.16 0.21 0.26	0.25 0.29 0.33	0.14 0.20 0.26	0.21 0.25 0.29	0.29 0.32 0.36	0.14 0.28 0.48
Hw reliability	**w.3.2.2**	0.42 0.45 0.50	0.43 0.45 0.48	0.26 0.29 0.33	0.36 0.37 0.38	0.43 0.48 0.53	0.48 0.53 0.59	0.38 0.42 0.45	0.42 0.45 0.50	0.29 0.33 0.38	0.33 0.36 0.38	0.26 0.42 0.59
Hw quality	**w.3.2.3**	0.18 0.23	0.23 0.25	0.48 0.51	0.36 0.37	0.05 0.10	0.21 0.26	0.25 0.29	0.30 0.34	0.38 0.42	0.29 0.32	0.05 0.31

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
		0.27	0.27	0.54	0.38	0.14	0.32	0.33	0.38	0.45	0.36	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 quality	**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
Co-determi-	**w.4.2**	0.00	0.14	0.16	0.01	0.18	0.17	0.13	0.06	0.03	0.00	0.00

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
nation		0.00 0.00	0.16 0.19	0.18 0.21	0.02 0.03	0.21 0.25	0.23 0.29	0.17 0.21	0.09 0.12	0.06 0.09	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 D4.2: Ranking of objectives and relevance from the online survey among ten stakeholders (see Tab. D1.1). "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
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

		Stakeholder										Overall		
Objective	**Weight**	**SH1**	**SH2**	**SH3**	**SH4**	**SH5**	**SH6**	**SH7**	**SH8**	**SH9**	**SH10**	**min**	**mean**	**max**
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 & manage-ment quality	w.4.1	**0**	**0**	**0**	**0**	4	1	**0**	**0**	2	**0**	0	0.7	4
Co-determi-nation	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 D4.3: Comparison of ranks and relevance of objectives in face-to-face interviews (Tab. D4.1) and the online survey (Tab. D4.2). 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 %)

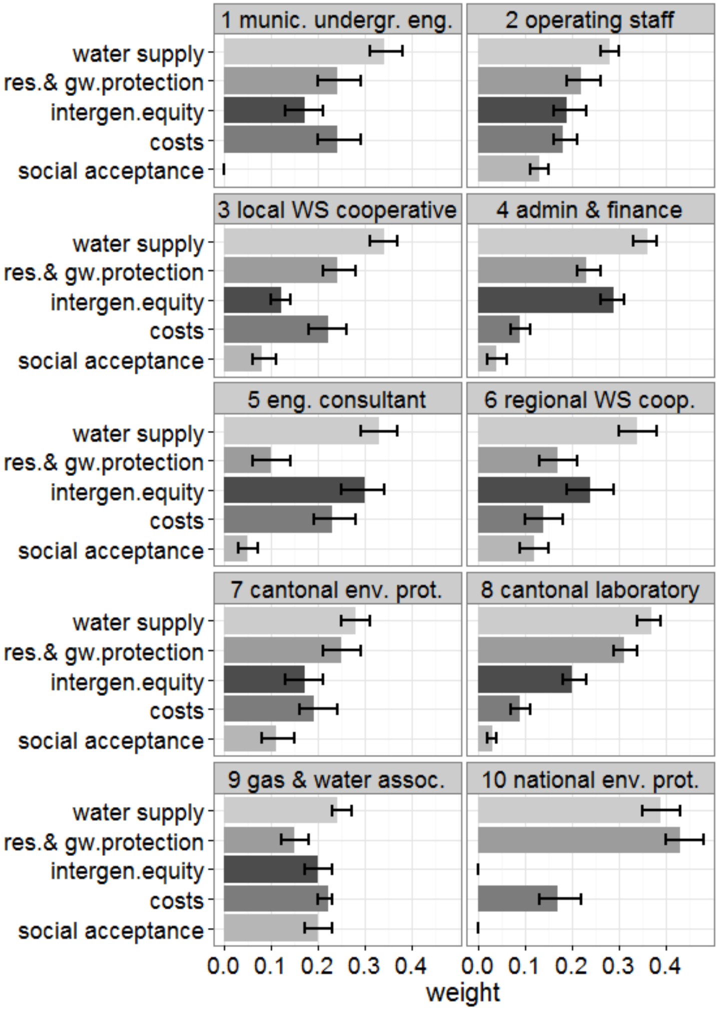


Figure D4.1: Weights of the sub-objectives of 'good water supply infrastructure', i.e. the five highest-level objectives for the ten stakeholders (see Tab. D1.1)

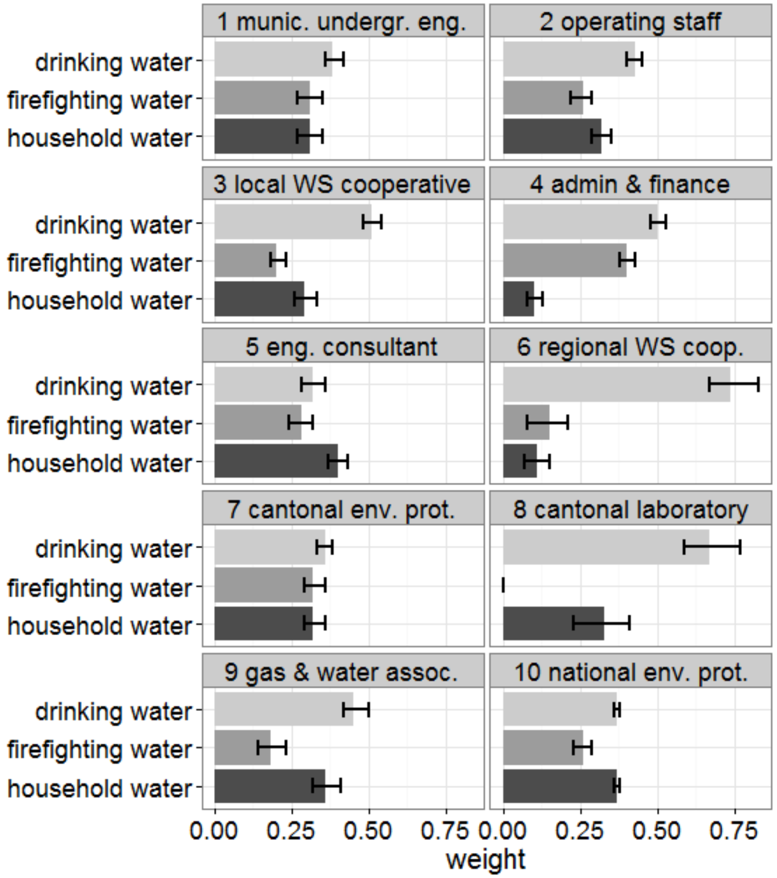


Figure D4.2: Weights of the sub-objectives of 'good water supply'

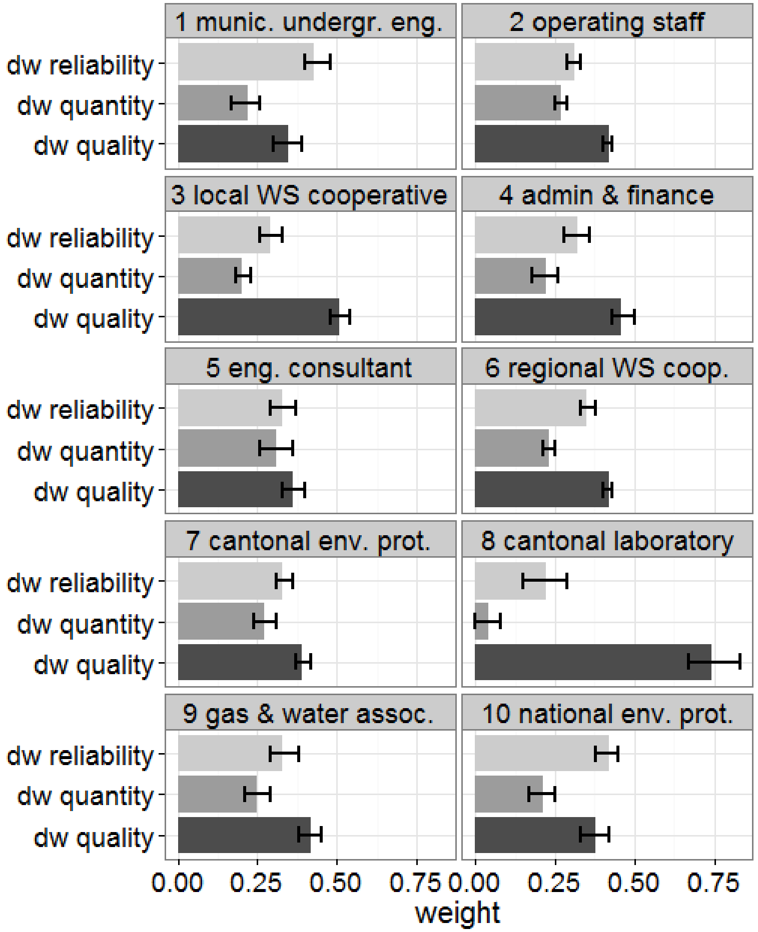


Figure D4.3: Weights of the sub-objectives of 'good water supply – drinking water'

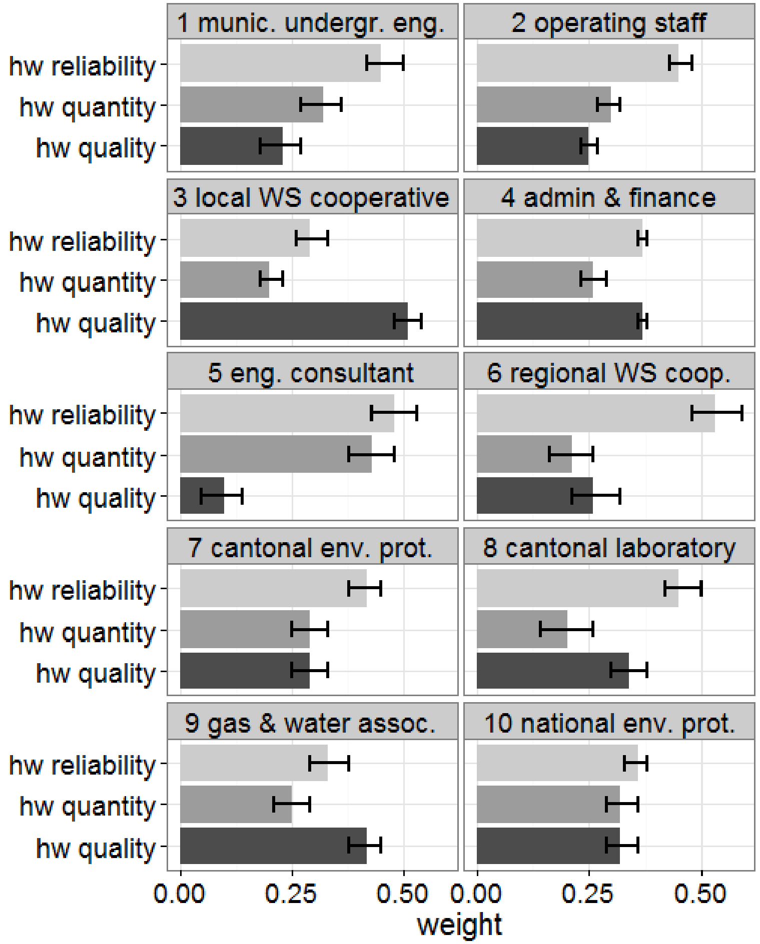


Figure D4.4: Weights of the sub-objectives of 'good water supply – household water'

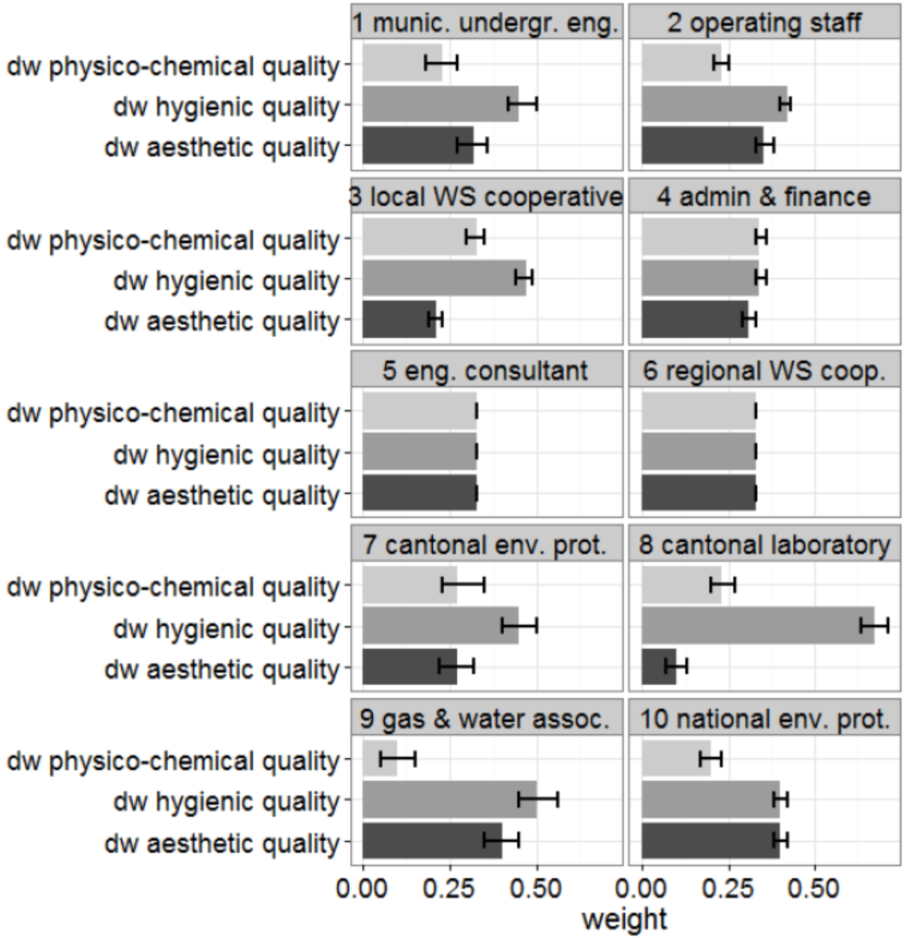


Figure D4.5: Weights of the sub-objectives of ‘ good water supply – drinking water-quality’

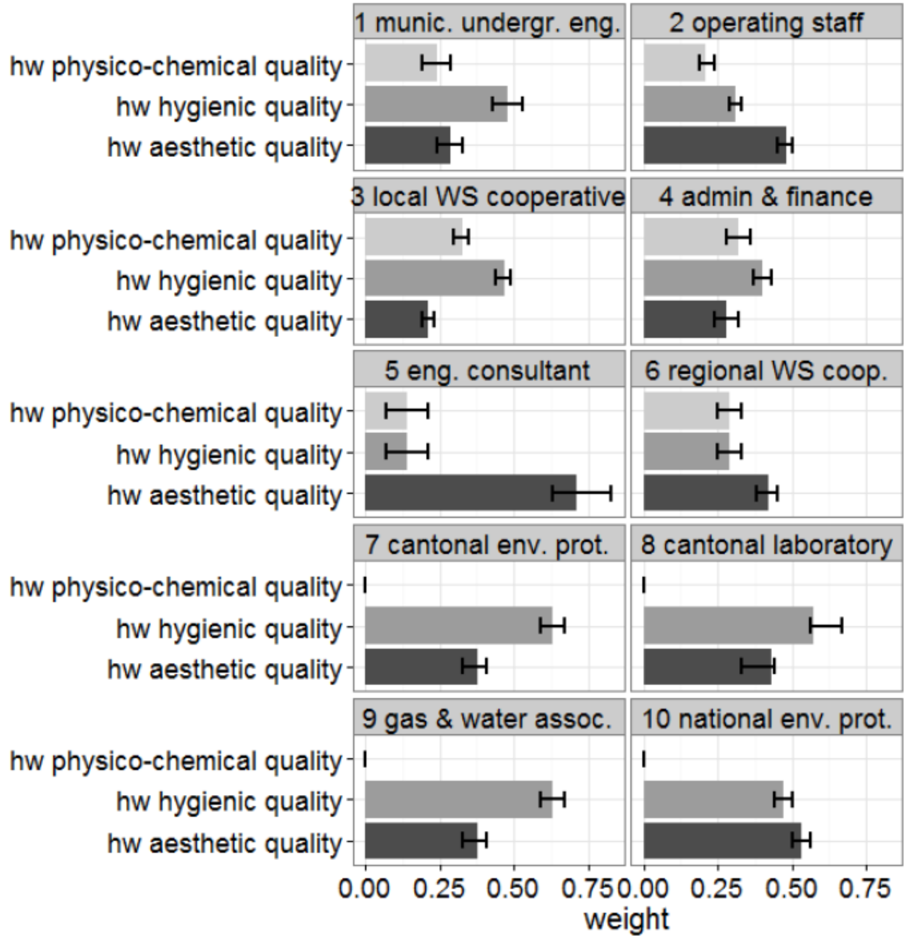


Figure D4.6: Weights of the sub-objectives of ‘good water supply – household water-quality’

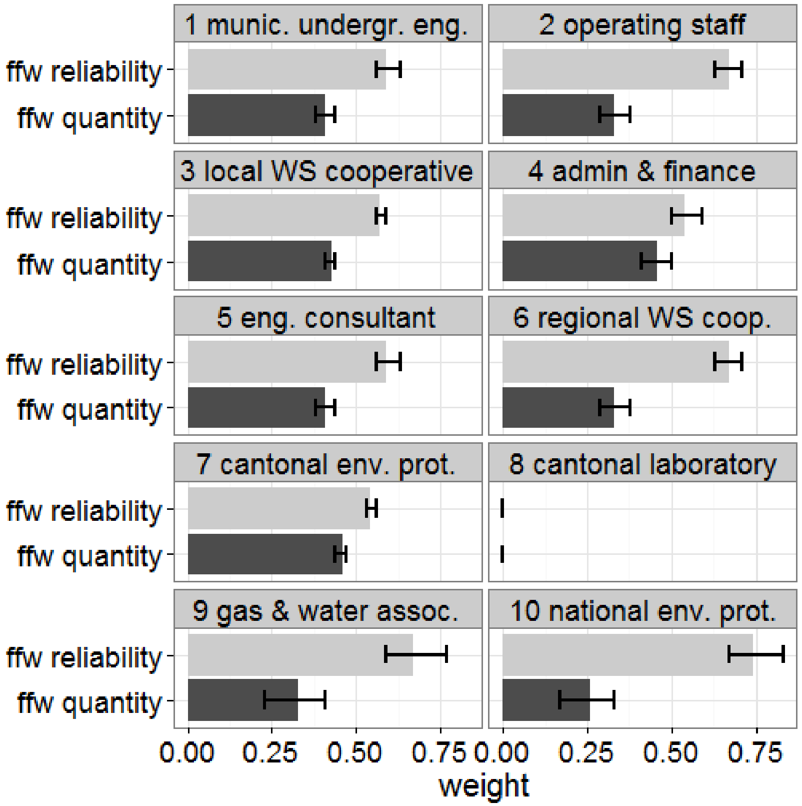


Figure D4.7: Weights of the sub-objectives of 'good water supply – firefighting water'

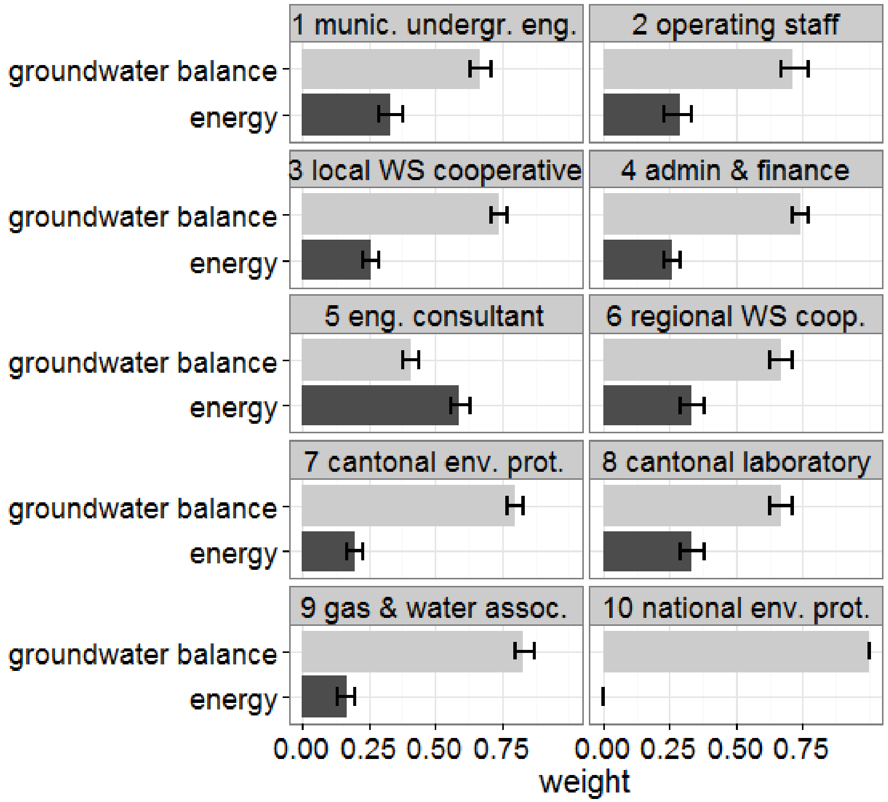


Figure D4.8: Weights of the sub-objectives of 'good resources and groundwater protection'

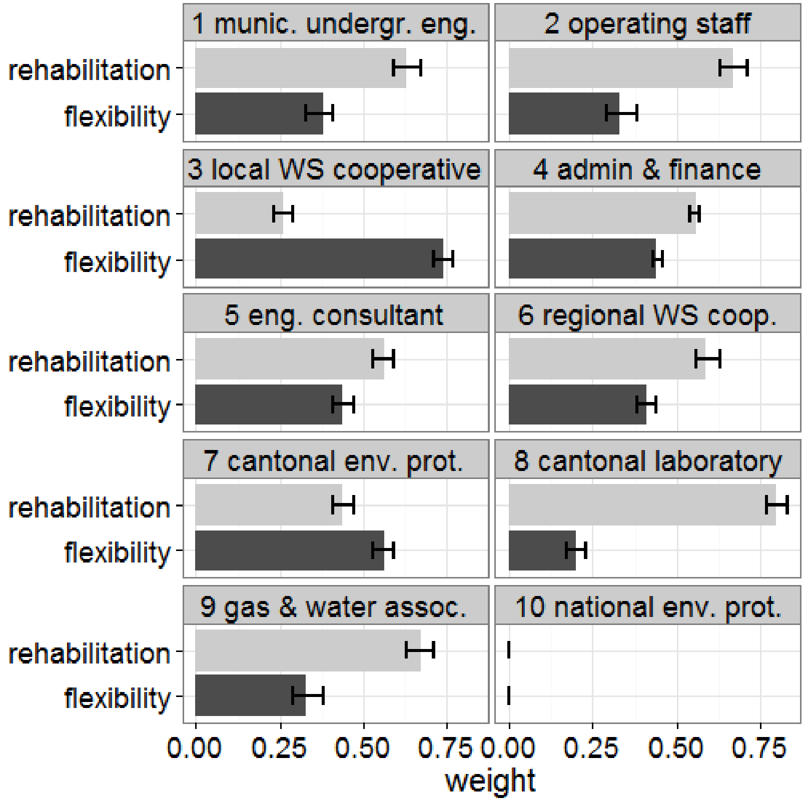


Figure D4.9: Weights of the sub-objectives of 'high intergenerational equity'

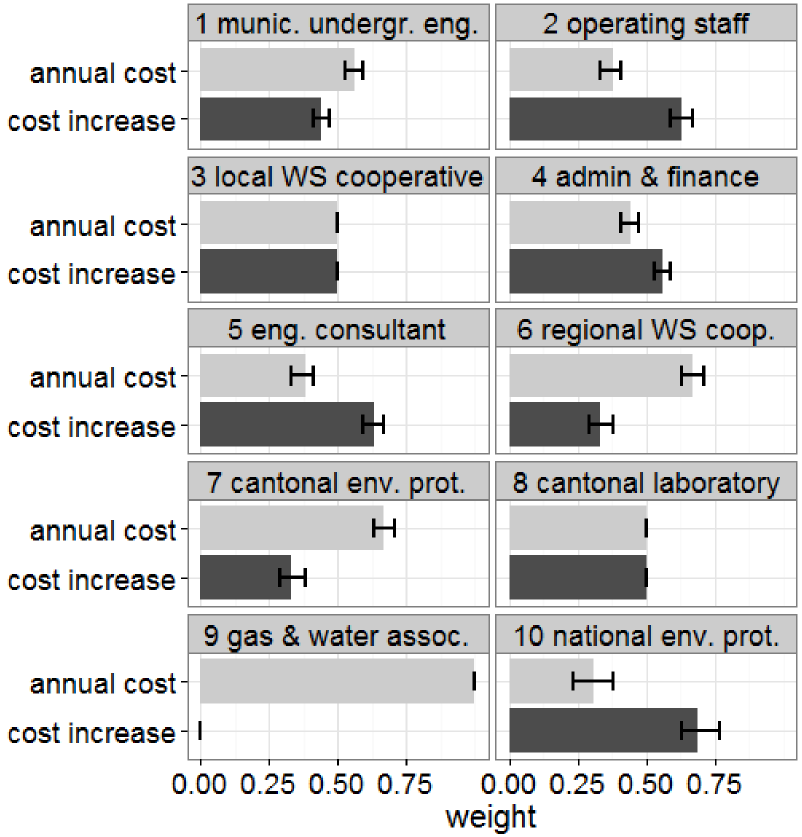


Figure D4.10: Weights of the sub-objectives of 'low costs'

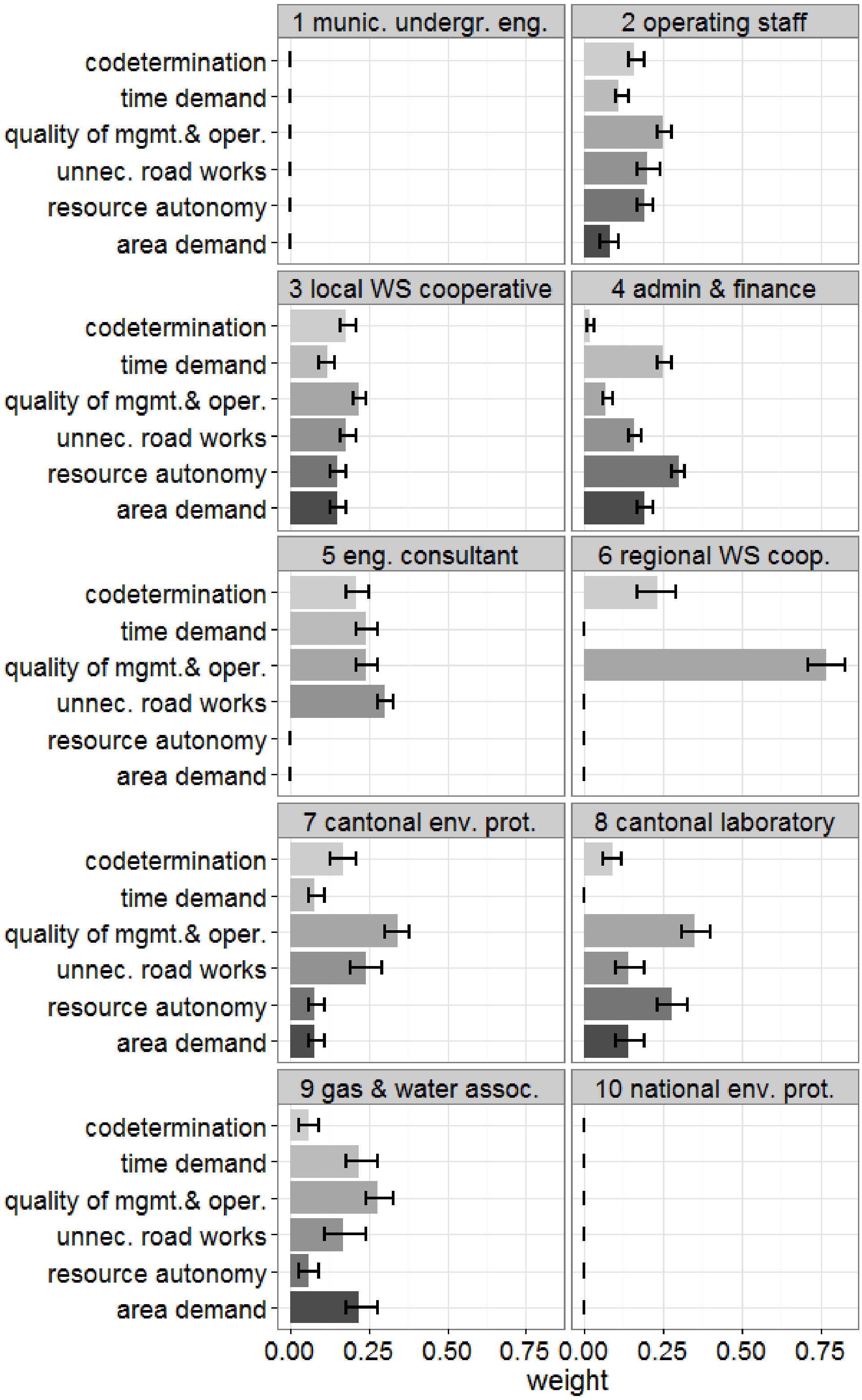


Figure D4.11: Weights of the sub-objectives of 'high social acceptance'

Marginal value functions

Table D4.4: 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_p, 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

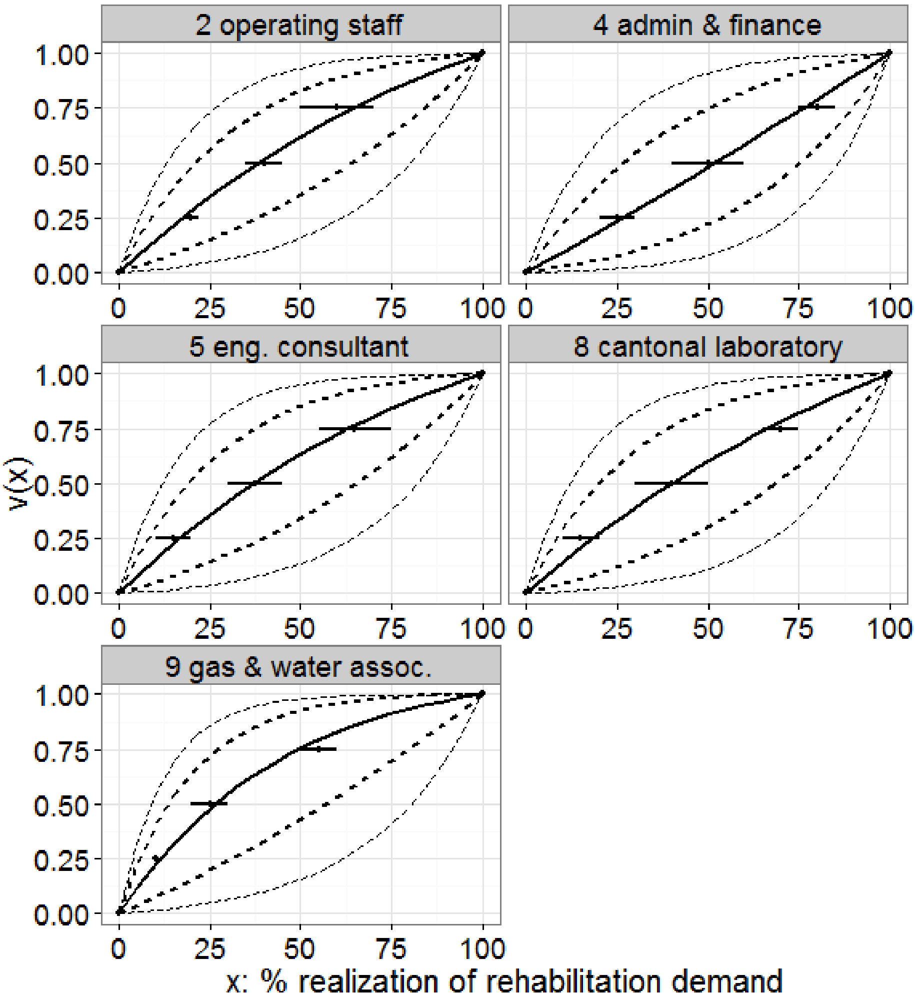


Figure D4.12a: Elicited value function levels for 'Realization of the rehabilitation demand [%]' (rehab) and fitted distributions for five stakeholders. The value v(x) = 0 on the y-axis indicates that this objective is not at all achieved, and 1 that it is fully achieved. Horizontal intervals show both 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. D4.4).

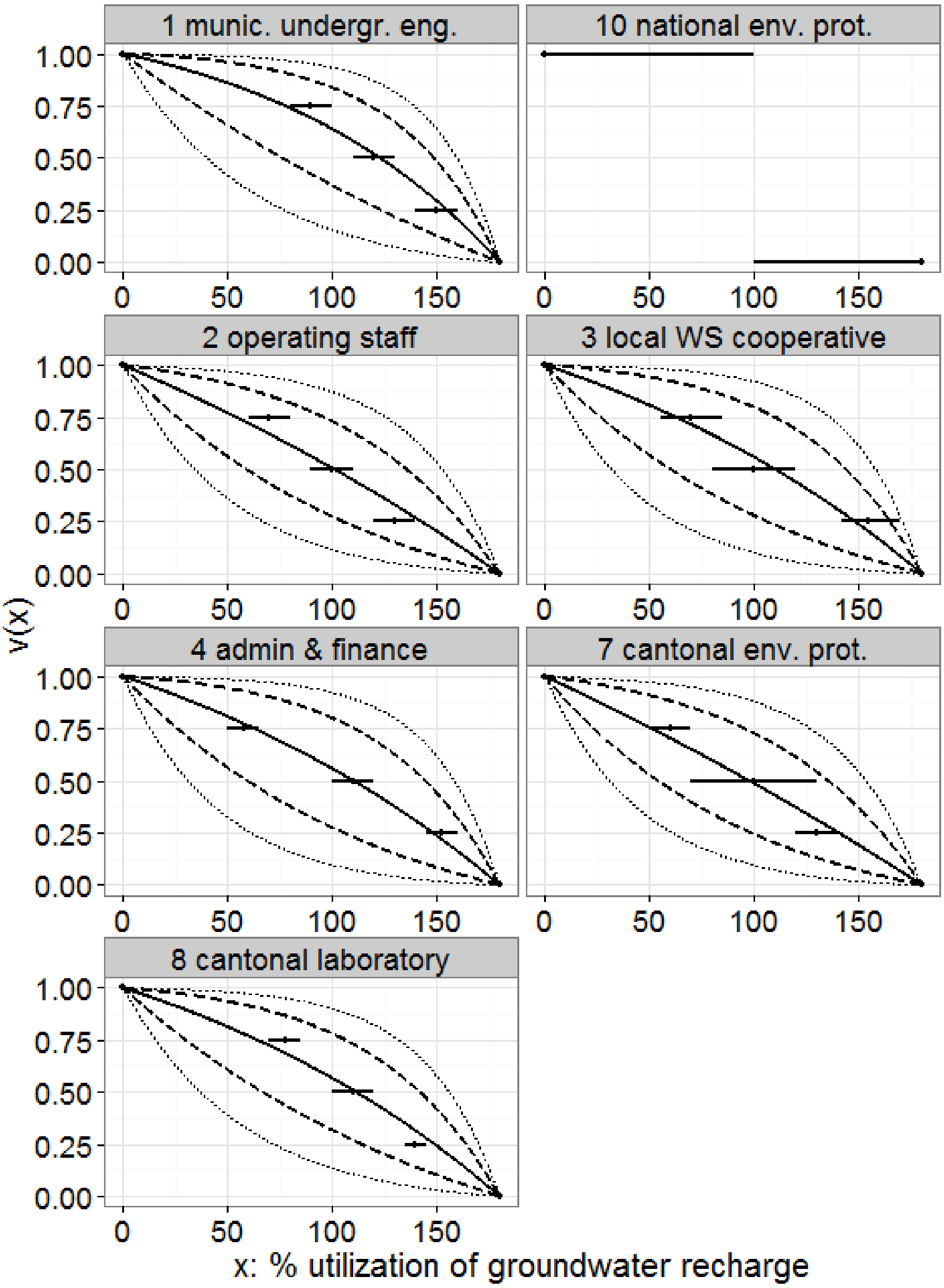


Figure D4.12b: Elicited value function levels for '% Utilization of groundwater recharge [%]' (gwhh) and fitted distributions for seven stakeholders. For detailed description see Fig. D4.12a.

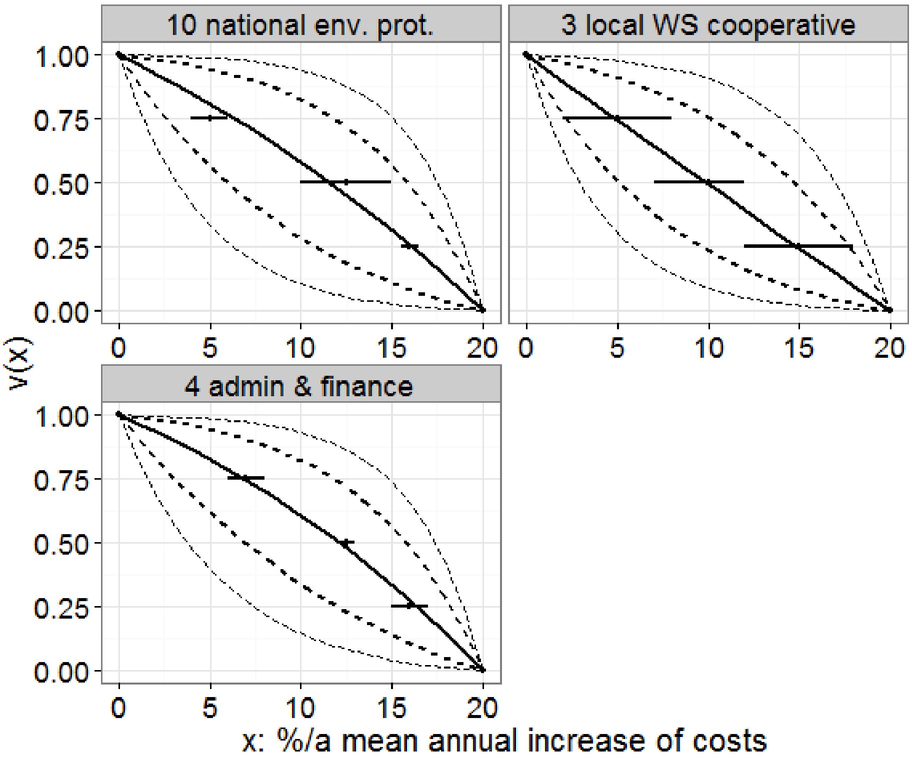


Figure D4.12c: Elicited value function levels for 'Mean annual (linear) increase of costs [%/a]' (costchange) and fitted distributions for three stakeholders. For detailed description see Fig. D4.12a.

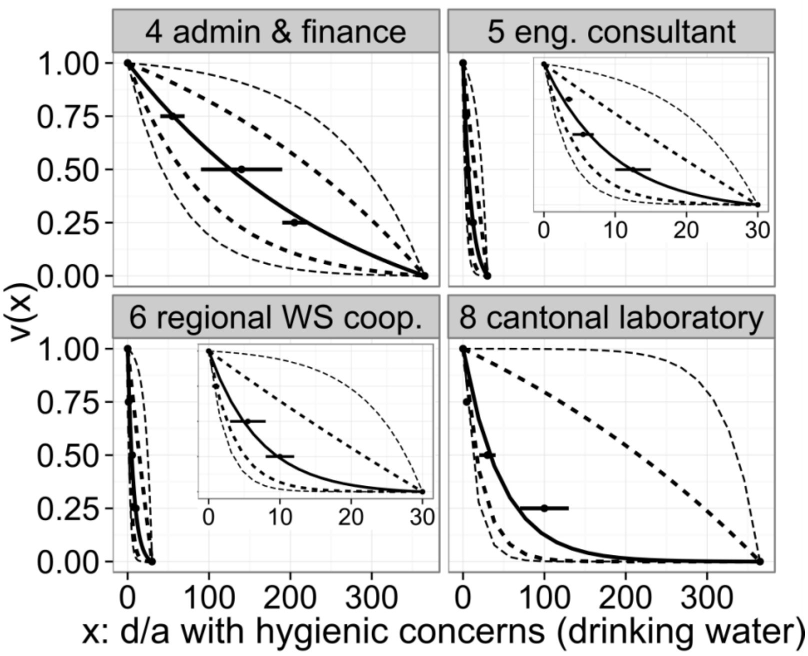


Figure D4.12d: Elicited value function levels for 'Days per year with hygienic concerns (drinking water)' (faecal_dw) and fitted distributions for four stakeholders. For detailed description see Fig. D4.12a.

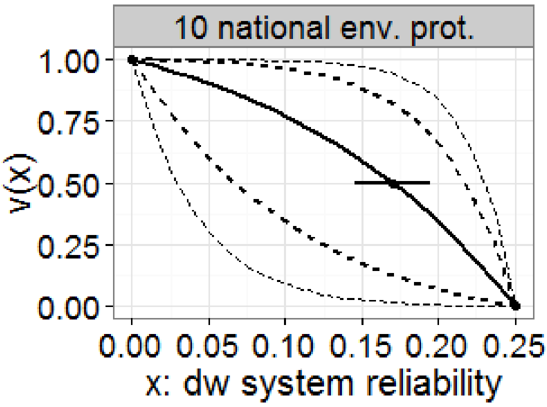


Figure D4.12e: Elicited value function instances for 'Drinking water system reliability' (reliab_dw) and fitted distributions used for uncertainty propagation for one stakeholder. For detailed description see Fig. D4.12a.

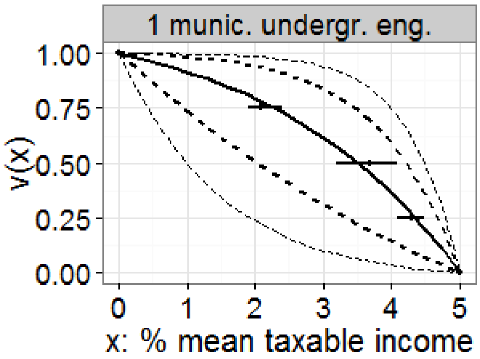


Figure D4.12f: Elicited value function levels for 'Annual cost per inhabitant in % of the mean taxable income [%] (costcap)' and fitted distributions for one stakeholder. For detailed description see Fig. D4.12a.

Marginal utility functions

Table D4.5: 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; $r\approx0$… 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 D4.6: 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.

D5 Uncertainty analysis

Table D5.1: Mean μ and standard deviations σ of rank distributions of 11 alternatives (A1a–A9; see Tab. D3.1) for four future scenarios and 10 stakeholders (SH; see Tab. D1.1)

	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 D5.2: Median rank (MR) and interquartile ranges (IQR) of rank distributions of 11 alternatives (A1a–A9; see Tab. D3.1) for four future scenarios and 10 stakeholders (SH; see Tab. D1.1)

	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 D5.3: 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**

D6 Global sensitivity analysis (GSA)

Stability of the ranking of alternatives to attribute sample size n

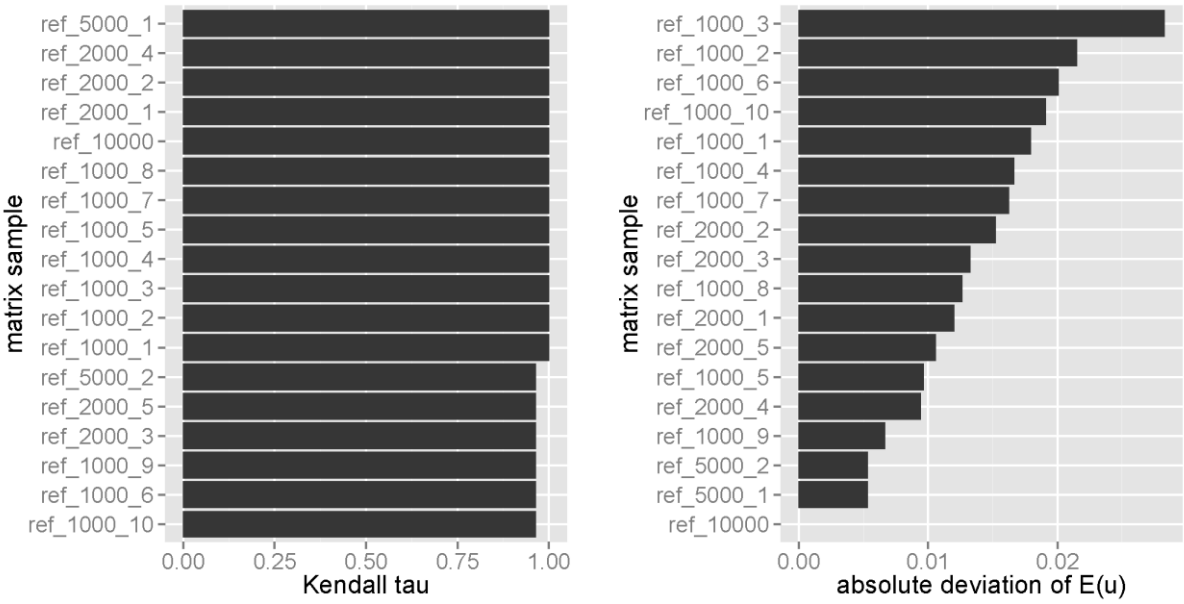


Figure D6.1: Change of Kendall rank correlation coefficient τ (left) and expected utility E(u) (right) depending on the size and portion of the underlying attribute sample. ref_10000 represents the reference using the overall n=10'000 attribute sample. The ranking using this sample was compared to sub-samples 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 D6.1: 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	

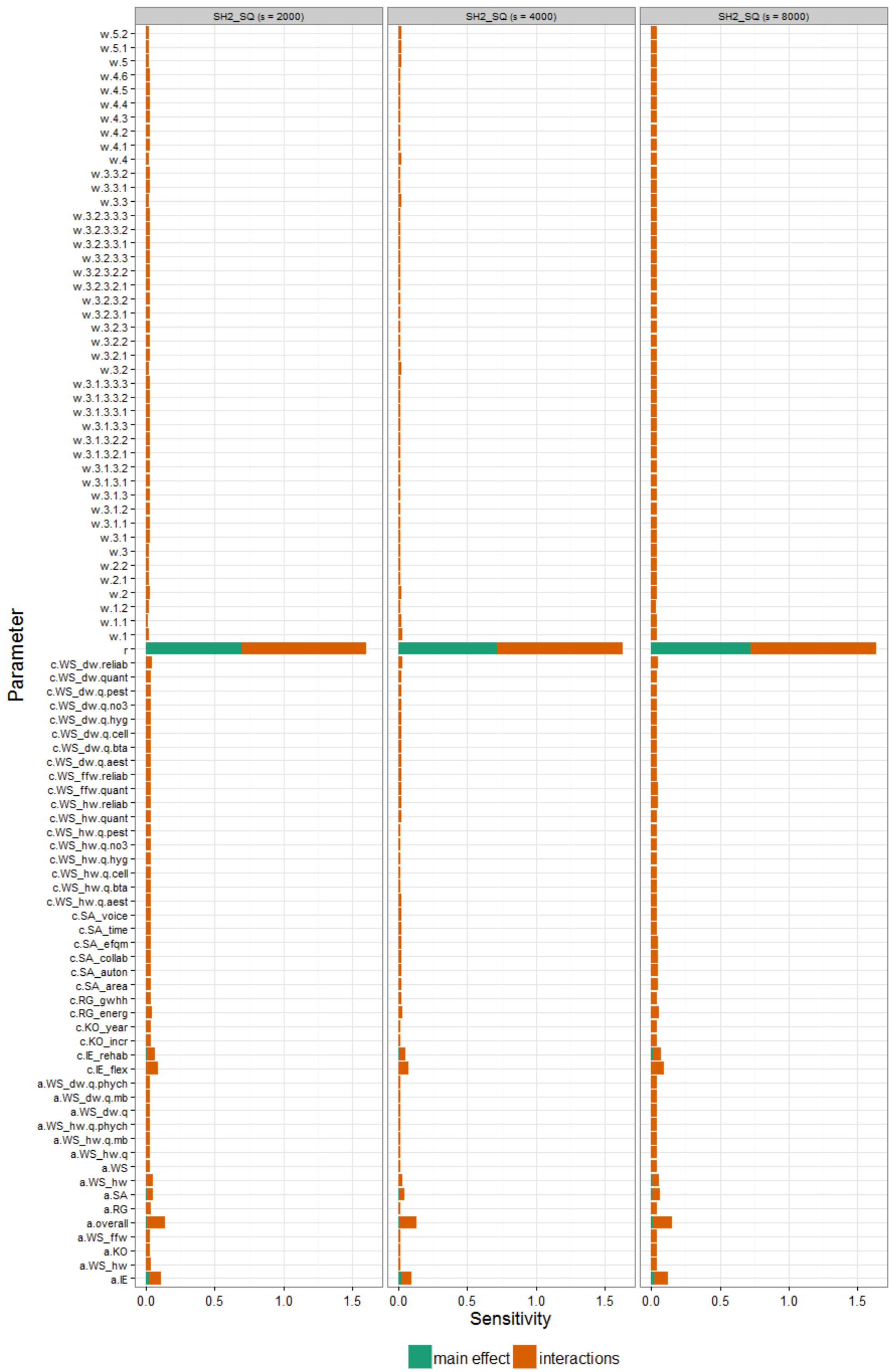


Figure D6.2: First and total order sensitivity coefficients for three different sample sizes s. Stakeholder SH2, Status quo scenario. r is the overall risk attitude, parameters starting with "a." are the aggregation mixture parameters, "c." value function curvature parameters, and "w." the weighting 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. 21) in main text), i.e. aggregation node or attribute. Acronyms for the top-level main objectives (and weight numbers, more details see Tab. D4.1) 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.

Curriculum Vitae

Education

2010- PhD student at EAWAG/ETH, Switzerland

Project: Sustainable Water Infrastructure Planning (SWIP)

2003-2009 Dipl.Ing. Water Resources Management (Wasserwirtschaft), TU Dresden, Germany

2005-2008 East Asia/China Studies, TU Dresden (Germany) and Nanjing University, PR China

Peer-reviewed publications

Scholten, L., Scheidegger, A., Reichert, P., Mauer, M., Lienert, J. 2013. Strategic rehabilitation planning of piped water networks using multi-criteria decision analysis. Water Research 49: 124-143.

Lienert, J, **Scholten, L.**, Egger, C., Reichert, P., and Maurer, M. 2013. Strategic decision making for sustainable water infrastructure planning under four future scenarios. Under review

Scheidegger, A., **Scholten, L.**, Maurer, M., Reichert, P. 2013. Extension of pipe failure models to consider the absence of data from replaced pipes. Water Research 47: 3696-3705.

Scholten, L., Scheidegger, A., Reichert, P., Maurer, M. 2013. Combining expert knowledge and local data for enhanced pipe survival modelling. Environmental Modelling & Software 42: 1-16