Accepted manuscript; European Journal of Operational Research

Methods to inform the development of concise objectives hierarchies in Multi-Criteria Decision Analysis

Mika Marttunen (corresponding author)
Swiss Federal Institute of Aquatic Science and Technology, Eawag
P.O. Box 611, CH-8600 Duebendorf, Switzerland
Finnish Environment Institute, SYKE
PB 140, 00251 Helsinki
mika.marttunen@environment.fi

Fridolin Haag
Swiss Federal Institute of Aquatic Science and Technology, Eawag
P.O. Box 611, CH-8600 Duebendorf, Switzerland
fridolin.haag@eawag.ch

Valerie Belton
University of Strathclyde, Business School
199 Cathedral Street
Glasgow G4OQU, Scotland, UK
val.belton@strath.ac.uk

Jyri Mustajoki
Finnish Environment Institute, SYKE
PB 140, 00251 Helsinki
jyri.mustajoki@environment.fi

Judit Lienert
Swiss Federal Institute of Aquatic Science and Technology, Eawag
P.O. Box 611, CH-8600 Duebendorf, Switzerland
judit.lienert@eawag.ch

Abstract

Building a well-structured objectives hierarchy is central to multi-criteria decision analysis (MCDA). However, in the absence of a systematic methodology to support the process, this task has been described as "more art than science". Objectives hierarchies often tend to become large and constraining the size of a hierarchy can be challenging. This paper proposes and illustrates the use of a set of methods to support the simplification of the hierarchies in contexts characterised by many objectives. Using data from two completed environmental cases we show retrospectively how qualitative (means-ends networks), semi-quantitative (relevancy analysis) and quantitative (correlation analysis, principal component analysis (PCA), local sensitivity analysis of weights) methods, used alone or in combination, can inform hierarchy development. We evaluate the potential benefits and challenges of each method and discuss the advantages and disadvantages of the simplification of an objectives hierarchy. Questionnaire-based relevancy analysis can be a useful method to identify and communicate important objectives in the early phases of an MCDA process with stakeholders, while correlation analysis can help to identify overlapping objectives, particularly in cases having many objectives and alternatives. It is intended that the methods support a facilitator in developing a clear understanding of the problem and also prompt deeper thinking about and discussion of the appropriate structure and content of an objectives hierarchy with the stakeholders involved.

Keywords: Problem structuring, Multiple criteria analysis, OR in environment, Behavioural OR

Highlights

- We propose a set of methods to support the development of objectives hierarchies.
- In particular, the methods can help to find ways to make a complex hierarchy more concise
- Data from two completed cases was used retrospectively to test and evaluate the methods.
- Relevancy analysis can be applied in the participatory screening of objectives.
- Correlation analyses help to identify overlapping objectives and critical trade-offs.

1. Introduction

Clear articulation of objectives can strongly improve the quality of a decision-making process (Keeney 1992). Identifying objectives and structuring them as a hierarchy is a central and perhaps the most challenging and time-consuming phase in Multi-Criteria Decision Analysis (MCDA) (Von Winterfeldt 1980, Bana e Costa and Beinat 2005). The size and structure of the hierarchy can profoundly impact later assessment phases and their outcome (e.g. Adelman et al. 1986), potentially influencing the weights allocated to the objectives and the results of MCDA (Borcherding and von Winterfeldt 1988, Weber et al. 1988, Weber and Borcherding 1993, Pöyhönen et al. 2001, Hämäläinen and Alaja 2008, Marttunen et al. 2018). Despite this, the development of objectives hierarchies has received surprisingly little attention in MCDA, (e.g. Adelman et al. 1986, Brownlow and Watson 1987). Keeney (1988) described this process as "much more art than science" suggesting that decision analysts simply have to develop appropriate skills (see also von Winterfeldt and Edwards 2007). However, whilst attention to problem structuring for MCDA has increased in recent years (Marttunen et al. 2017), Brugha (2004, 2015) suggests that there is still a lack of sound procedures for developing objectives hierarchies that are suitable for use in formal MCDA.

In the 1980s, several studies dealt with the hierarchy development process, considering top-down and bottom-up approaches[1] (Adelman et al. 1986, Buede 1986). Brownlow and Watson (1987) raise two central questions: would different ways of approaching structuring the hierarchy have led to different hierarchies and when should one stop modifying hierarchies? Thirty years later, these questions are still valid. In the late 1980s and 1990s, research interest turned from the hierarchy building process to its structure, demonstrating how variations in the hierarchy structure can affect the allocation of weights to objectives (Borcherding and von Winterfeldt 1988, Pöyhönen and Hämäläinen 1998). These studies on the "splitting bias" showed that splitting or merging an objective can increase or decrease its weight even though the information content of the objectives remains the same (Weber et al. 1988, Pöyhönen et al. 2001).

More recently, research about hierarchies has broadened. Topics have included the generation of objectives using value-focused and alternative-focused thinking (Bond et al. 2008, 2010, Selart and Johansen 2011) and using problem structuring methods to support hierarchy building (Montibeller and Belton 2006). Automatic or unaided methods for hierarchy structuring emerged as a new research topic;

[1] Top-down: upper-level objectives are listed first then each objective is divided into sub-objectives. Bottom-up: start from the lowest-level objectives without considering the hierarchy structure; later, cluster lowest-level objectives to form the hierarchy branches.

e.g. semantic analysis via statistical web mining (Maida et al. 2012), or an automatic clustering procedure (Maier and Stix 2013). Increased interest in behavioural operational research (Hämäläinen et al. 2013, Montibeller and von Winterfeldt 2015, Franco and Hämäläinen 2016) has also given rise to new research about objectives hierarchies (Marttunen et al. 2018). However, there is little research which explores systematic approaches to the identification of relevant objectives. We identified only two papers which do so (Braunschweig et al. 2001, Srdjevic et al. 2012). The aim of this research was to explore ways of broadening and strengthening this process.

The focus of this paper is on the potential to support the decision analyst / facilitator in the process of helping the decision makers create an appropriate objectives hierarchy in contexts which are "data rich". Examples include: procurement (in particular public procurement) in which the procuring organisation is the recipient of very detailed invitations to tender; and environmental decision making, which is characterised by a wide range of stakeholder concerns and extensive monitoring data or impact predictions. Such situations often lead to a desire on the part of those tasked with making the decision to make sure "everything is taken into account". This can lead to very large criteria hierarchies which are more "descriptive" of the situation rather than reflecting what really matters in the specific context (see Marttunen et al. 2018).

We describe a systematic procedure and associated methods to inform decisions regarding the size and structure of the hierarchy in a specific situation where a list of candidate objectives, a set of alternatives and some data on their performance with respect to the objectives is available. Each of the proposed methods has been used, independently, by one of the five authors of this paper in previous MCDA interventions. Our aim is to illustrate how they can be used effectively, either independently or in conjunction with each other, to support a facilitator in helping decision makers define a concise objectives hierarchy which has a good balance between completeness, in that it includes the central objectives, and simplicity, in that the elicitation burden is not excessive. In essence, it should satisfy Phillips' (1984) concept of a requisite decision model. The potential use and value of the proposed methods are illustrated through their retrospective application and evaluation in the context of two completed case studies (each of which involved a different author of this paper). The analyses presented here are based on the original data from these two real-world, environmental decisions, both of which involved engagement with multiple stakeholder groups[2].

[2] Note that no stakeholders in the original cases were involved in the retrospective analyses. Two authors of this paper were each involved as a facilitator in one of the original cases.

This paper complements our earlier work associated with a two-year research programme related to problem structuring and model building for MCDA. An initial literature review explored the use and associated strengths and weaknesses of specific methods (e.g. SWOT, Stakeholder Analysis) and broader methodologies (e.g. SODA, SSM) in supporting problem structuring for MCDA (Marttunen et al. 2017). A second paper focused on the influence of characteristics of an objectives hierarchy on the objectives weights (Marttunen et al. 2018).

2 Constructing objectives hierarchies

2.1 Phases in hierarchy building

The first stage in any decision process is the understanding of the broader context and characteristics of the problem (Fig. 1). This includes the identification of stakeholders and their interests, concerns and hopes. During this diverging or opening-up phase a comprehensive representation of the problem is developed which may be presented, for example, in the form of a cognitive or causal map, a means end network or an objectives hierarchy. MCDA is appropriate for well-structured problems; it does not, per se, include any specific tools for the problem structuring phase although nowadays, it is increasingly applied in concert with problem structuring methods (Marttunen et al. 2017).

In this paper, we consider a situation where the initial problem structuring phase has led to such a large objectives hierarchy that it is considered necessary or appropriate to explore possibilities to simplify it for preference modelling. Hence, we focus on potential methods for reducing the size of the objectives hierarchy. It should be noted that some structuring processes may start from a simple objectives hierarchy and develop it step by step towards a more comprehensive and potentially more complex one as new objectives emerge. Readers interested in the approaches to support the diverging phase are referred to Keeney (1992), Rosenhead and Mingers (2001), French et al. (2009) and Belton and Stewart (2010).

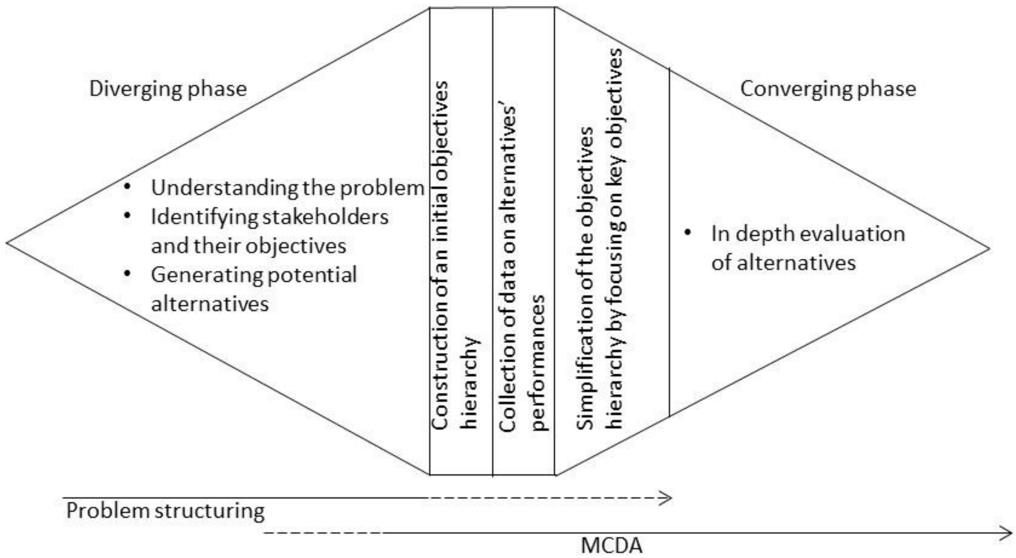

Figure 1. Diverging and converging phases of decision-making and stages in building an objectives hierarchy.

2.2 Challenge in hierarchy building: large objectives hierarchies

Objectives hierarchies often tend to become large and constraining the size of a hierarchy can be challenging. This is common in many domains, such as environmental, energy, transport and military decisions and public procurement (see Marttunen et al. 2018; and for examples of large hierarchies see Keeney et al. 1987, Langhans et al. 2014, Rahman et al. 2015). The main reasons why objectives hierarchies often become large in environmental decision situations relate both to the complexity of the decisions and to factors motivating different stakeholders and the facilitator, for example:

- Many economic, social and ecological impacts need to be assessed to cover all impact dimensions and legal requirements (e.g. Kiker et al. 2005).
- Various stakeholders with different interests are engaged. They know their own field very well and specify it in detail when objectives are identified (von Winterfeldt and Fasolo 2009).
- Ecologists and engineers use a large number of indicators and may wish to include all of these in the objectives (Lienert et al. 2015, Langhans and Lienert 2016).
- Stakeholders wish to include objectives that are likely to favour their preferred alternative (von Winterfeldt and Fasolo 2009).
- Decision analysts may aim to satisfy the concerns of stakeholders and appreciate their opinions even if exclusion of certain objectives could be justified (von Winterfeldt and Fasolo 2009).
- The aim to minimise post-decision regret can result in very detailed analysis and excessive information requests (Retief et al. 2013).

This observation that hierarchies tend to grow too much appears to be contradictory to the finding that people may have difficulties in recognizing their objectives (Bond et al. 2008). However, there are some important differences between the two contexts. The study by Bond et al. (2008) relates to unaided personal decisions (which are hypothetical) in an experimental setting, whilst the focus here is on large, complicated public policy decisions, which are usually guided by an experienced facilitator who aims to include the interests of a diverse set of stakeholders. Moreover, in environmental cases there are many existing indicators from natural science, which stakeholders want to see included as objectives in the hierarchy.

MCDA processes can produce such a large amount of information concerning stakeholders' preferences, alternatives' impacts and their overall values that it exceeds the cognitive capacity of participants. This "information pollution" (Hobbs and Meier 2000) is a potential weakness of MCDA if structuring is not done properly (Bana e Costa and Beinat 2005). It is crucial to find a good balance between completeness and conciseness in an objectives hierarchy. After a certain point, adding more detail to the hierarchy can divert attention from the most relevant objectives, it can increase biases in weight elicitation and lead to improperly laborious processes (see e.g. Adelman et al. 1986, Kleinmuntz 1988, Hobbs and Meier 2000, von Winterfeldt and Edwards 2007). On the other hand, reducing an objectives hierarchy may result in losing information which can help to discriminate between alternatives and objectives may be open to misinterpretation because they have become very general and encompass too many aspects that are not explicit (e.g. Mathieson 2001).

2.3 Characteristics of good objectives and hierarchies

A good objectives hierarchy has an appropriate content and structure, taking into account the requirements of MCDA. Characteristics of a good set of objectives that constitute the hierarchy have been presented in several MCDA textbooks (Keeney and Raiffa 1993, Belton and Stewart 2002, Eisenführ et al. 2010, Clemen and Reilly 2014). Although the terminologies vary, their essence is the same:

- Completeness: all important consequences of alternatives and stakeholders concerns in a decision context can be adequately described with the objectives.
- Conciseness: being brief in form but comprehensive in scope.
- Non-redundancy: no double counting or partially overlapping objectives.
- Understandability: the objectives can be understood by any interested individual.
- Measurability: it should be possible to define attributes which measure the achievement of objectives to a level appropriate for the analysis. This could imply the use of a natural, proxy or a constructed scale (Keeney and Gregory 2005).
- Preferential independence: if an additive MCDA aggregation model is used, trade-offs between any two objectives are independent of the outcomes on other objectives.

We also suggest that MCDA modellers should give careful consideration to the structure of the hierarchy (i.e. number of hierarchy levels and number of lowest level objectives in each hierarchy branch) in order to avoid the range insensitivity bias and potential asymmetry effect which can occur when using hierarchical weighting (Marttunen et al. 2018).

Brugha (2004) incorporates the above characteristics for a good objectives hierarchy in guidelines for the broader MCDA process, including the potential later use of a model, and Goossens et al. (2015) discuss the use of this framework in practice. Brugha (2004) and O'Brien and Brugha (2010) add to the above considerations, stating that: information relating to objectives should be *accessible*; the objectives hierarchy should be *abstractable* (objectives can be interpreted in a broader context); *verifiable* (decision makers own and can justify all aspects of the hierarchy); *refinable* (decision makers have the opportunity to adjust alternatives and scores, objectives and weights, set of alternatives); and *usable* (it provides preferences without excessive effort or cost). This last criterion is one of the key motivations for our research. Keeney and Gregory (2005) also emphasise the selection of attributes (the selected means of measuring performance against objectives) as inadequate ones can diminish the value of an analysis which is based on a well-structured hierarchy. The desirable properties of attributes they propose are similar to the characteristics of a good hierarchy presented in the bullet points above.

2.4 Structuring and positioning individual objectives

Usually there are many potentially good representations of the decision problem, which may differ in terms of hierarchy size and structure. Objectives can be defined and labelled in different ways and they can be split, merged or located differently in the hierarchy. Consideration of the following operations can stimulate alternative ways of building a hierarchy:

Operations which increase the number of objectives:

- Adding an objective which provides new information about the alternatives' performances.
- Dividing an objective into two (or more) objectives at the same hierarchy level.
- Describing an objective in more detail by adding sub-objectives, creating a new hierarchy level.

Operations which decrease the number of objectives:

- Removing an objective and thus its information about alternatives' performances.
- Removing a set of sub-objectives leaving the higher-level objective to capture all the associated information.
- Combining two (or more) objectives at the same hierarchy level into a new objective which captures all the information.

Operations which change the location of an objective:

- Transferring an objective to a higher or lower level of the hierarchy.
- Transferring an objective from one branch to another.

Typically, decision-makers are interested in the trade-offs at a high level, for example between costs and benefits or between ecological, social and economic impacts, to which it might be difficult to assign weights. These trade-offs may be more clearly highlighted and decision-makers' thinking informed by a simple visualisation. For example, if there are only two main objectives the results from each branch can be presented in an x-y graph. The results of cases having three main objectives can be shown as a triangle (Hodgkin et al. 2005) and four or more as a line graph/value path or radar chart (e.g. Valle and Climaco 2015). This is something a decision analyst may wish to take into account when structuring the objectives hierarchy.

2.5 Pros and cons of simplifying the hierarchy

The trade-off between being concise and complete can be challenging. There is no unequivocal criterion for determining when a more extensive hierarchy structure is better than a concise one (León 1999). Keeney and Raiffa (1993, p. 43) suggest that the advantages and disadvantages of further specification should be analysed. Potential positive and negative consequences of simplification are summarised in Table 1. In general, the larger the initial hierarchy, the greater are the potential advantages and the smaller the potential disadvantages of the simplification. It is important that the analyst is aware of these consequences and can ensure an informed discussion and choices regarding the size and structure of the hierarchy, taking into account its anticipated uses. For instance, a stakeholder workshop can be organised where a candidate objectives hierarchy is systematically evaluated or two hierarchy candidates are compared with respect to the issues in Table 1. In some cases, it can be desirable to keep technically irrelevant objectives in the analysis because this easily allows stakeholders to see that these objectives were considered in the decision process.

Table 1. The potential advantages and disadvantages of simplifying objectives hierarchies.

Potential advantages	Potential disadvantages
Interaction and communication	
• Most important trade-offs are clearly visible • The hierarchy is easier to grasp • More in-depth discussion of each objective if time is limited	• Loss of important information • Stakeholders might not accept exclusion of objectives • New, merged objectives cover too many aspects that cannot be easily measured as a whole • Less transparent, making it more difficult to communicate to external parties
Efficiency	
• Less effort for assessing impacts, scoring and preference (e.g. weight) elicitation	• Scoring of alternatives is challenging for merged objectives • Preference elicitation (e.g. weighting) is more challenging as objectives have a broader definition and more dimensions have to be considered
Risk of cognitive biases	
• Smaller cognitive load in e.g. weight elicitation • Reduced risk of overriding important objectives by a large number of less important objectives	• Increased risk of inverse splitting bias in weight elicitation as objectives are merged (i.e. weight of merged objective is lower than if two objectives are presented separately)

3. Methods to support the building of concise hierarchies

Various qualitative and quantitative methods can improve the decision analyst's understanding of the relevancy of the objectives and how they relate to each other. We illustrate how means-ends networks, relevancy analysis, statistical analyses (correlation analysis, principal component analysis) and sensitivity analyses can inform decision analysts in hierarchy building and evaluate the applicability of these methods in general.

The information needs of each method determine the appropriate point of use (Fig. 2). For instance, relevancy analysis can be conducted in the early phases of MCDA, as a part of problem structuring to screen out the least relevant objectives prior to impact assessment. After the impact assessment phase has been realised, it is also possible to do (or even redo) the relevancy analysis informed by the impact predictions. Statistical analyses, which require information about the alternatives' performance scores, can typically only be applied later when that information is available Sensitivity analyses are usually performed at the end of an MCDA, but we propose that they may also be performed earlier to inform the decision about excluding objectives.

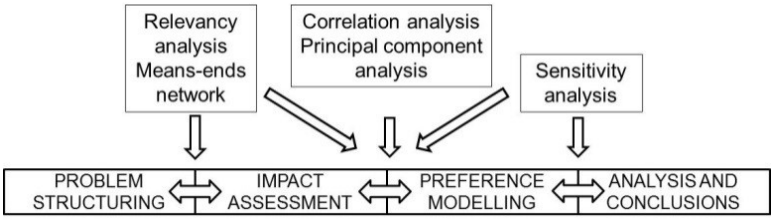

Figure 2. Potential to apply the proposed methods in the MCDA process.

3.1 Means-ends networks

Means-ends networks distinguish fundamental objectives, the ends we are trying to achieve, and means objectives, which are important in achieving these ends. There is a long tradition of using means-ends networks in the building phase of objectives hierarchies (Gregory et al. 2012). Keeney (1992) proposed their use to show relationships among actions and to highlight that for many of the valued ends, different means could be used to achieve them. More knowledge about these relationships can help to identify complementarities and redundancies across objectives (Gregory and Wellman 2001).

3.2 Relevancy analysis

Keeney and Raiffa (1993, p. 43) suggested a "test of importance" to decide whether an objective should be included in the hierarchy. Hereby, the decision maker is asked *"whether he feels the best course of action could be altered if that objective were excluded".* To our knowledge, this test is not widely used in MCDA, presumably because answering the question in one's head is cognitively very demanding and prone to error. Motivated by this concept, we developed a systematic analytic procedure, which we call relevancy analysis, to support experts and stakeholders in reaching an agreement about which objectives need to be included in the MCDA.

Relevancy analysis can be classified as a qualitative or quantitative approach depending on the way it is applied. In this paper, we demonstrate the use of approach relatively early in the problem structuring phase of an intervention to identify which of the initially identified objectives are most important to include in the model (because they are judged to be both highly "important" and also significantly discriminate between the alternatives). However, once developed it can also inform the specification of objectives' weights. Marttunen et al. (2019) demonstrate the quantitative use of relevancy analysis in a case where relevancy analysis was applied in the identification of the main expected benefits of one river restoration measure in Switzerland.

Relevancy analysis is conceptually similar to the swing weight matrix method (Ewing et al. 2006, Parnell and Trainor 2009, Parnell et al. 2013). However, the swing-weight matrix is a quantitative approach developed specifically for use at the later stage of assigning quantitative values to objectives' weights. It explicitly considers both the importance and the range of the "value measures" in assigning weights to objectives. The swing weight matrix method consists of five steps: (i) define the importance and variance dimensions; (ii) place the value measures (attributes, we use the term objectives in our cases) in the matrix; (iii) assess the unnormalised swing weights (0-100) of the attributes; (iv) calculate the overall weights; and (v) explain the normalised swing weights (Ewing et al. 2006, Parnell and Trainor 2009). However, the methods were developed with different aims. Relevancy analysis as proposed here seeks to provide a systematic support for the identification of key objectives for inclusion in the objectives hierarchy at the early phase of MCDA. This could potentially be before data relating to the alternatives is available or after impact assessment (see Fig. 2)[3]. The swing weight matrix method, on the other hand, is a way to elicit and present the weights for objectives in a consistent and understandable way when impacts of alternatives are known (Ewing et al. 2006, Parnell and Trainor 2009).

Relevancy analysis also has similarities with the approach applied in environmental impact assessment procedures (EIA) to assess the significance of impacts (e.g. Wood 2008, Terrapon-Pfaff 2017). The significance of an impact in EIA is determined as a combination of the importance of the objective in general and the magnitude of its potential impact in the specific application case (Wood 2008). The aim is to identify those impacts that need most attention in the assessment process, in the comparison of alternatives and in the design of measures to mitigate adverse impacts of the project.

Relevancy analysis brings principles from impact significance assessment to the MCDA context. In common with the swing weight matrix method, it separates the two conceptual dimensions of an

[3] As our analyses of the case studies were retrospective, the impact data was available and thus the impact ranges were established.

objective's weight in an MCDA based on multi-attribute value theory (MAVT; Keeney and Raiffa 1976), namely the "general importance" of the associated concern and the extent to which the objective differentiates between alternatives (problem specific impact range). As part of this study, we focused on finding good questions in order to operationalise relevancy analysis. An example of the questions used in an environmental case is presented in Annex 1. The estimation of the importance of the objective is based on questions 1 to 4 (below) and the problem specific impact range is estimated with one question (5). Similar questions are typically used in environmental impact assessments (e.g. Lawrence 2007, Wood 2008).

1. Are there specific nature, society or economy related values with respect to the objective and what is their importance? (e.g. if a recreational area has been designated as regionally or nationally significant, it increases the importance of the recreation objective). Note: Values which are associated with the other objectives in the hierarchy should not be considered to avoid double-counting.
2. Are there international, national, regional or municipality level legislation or recommendations related to the objective and how strict are they (scale: no, non-binding regulations, binding regulations)?
3. How easily is the objective affected or how well does it tolerate changes (e.g. are small changes significant, are there any particularly sensitive targets/objects in the study area)?
4. Bearing in mind the responses to questions 1–3: how important is it to (i) improve the achievement of the objective or (ii) prevent the deterioration of the objective from the current state in the study area (scale: not important, slightly, moderately, important, very important)?
5. How large is the difference between the worst and best cases with regard to the objective? When giving your estimate also take into account if there are differences in the spatial extent or duration of the alternatives' impacts (scale: low, moderate, high, very high).

Tentative guidance for answering the importance assessment questions (Question 4) is given in Appendix 1[4]. The results are presented as a two-dimensional chart which can support the identification of objectives to be included in MCDA (see S-1-1 and also Fig. 5 in the main text). For instance, objectives which belong to the "Low importance" and "Low impact range" category might be excluded, or an objective might be a candidate for exclusion if either its importance or impact range is classified as "Low".

3.3 Statistical analyses

[4] Two subject experts, a consultant (LUP case) and a water management engineer (SWIP case) responded to a relevancy analysis questions relating to their area of expertise.

3.3.1 Correlation analysis and principal component analysis

If data on the impact/ performance of alternatives is available (even in a preliminary form as rough estimates), this multi-variate dataset can be explored visually (e.g. scatterplots) or with statistical tools. This can give insights into the structure of the decision problem and relationships between the objectives that may be helpful in the structuring process. We describe two potentially useful methods, correlation analysis (CA) and principal component analysis (PCA), acknowledging that other ways to explore such data exist (e.g. factor analysis; Slovic et al. 1986). Both methods can inform MCDA in many ways, but to our knowledge have not been used to assist hierarchy building. To be able to analyse the alternatives' performance on value scales (0–1), assumptions about the lowest-level value functions are required[5]. In this study, we assume linear marginal value functions.

Correlation analysis (Cohen and Cohen 1983) helps to detect relationships between the alternatives' performances (impacts) on different objectives. We use it to identify potentially redundant objectives, which might be eliminated or combined. For variables on a continuous scale, Pearson's correlation coefficient measures the strength of their linear relationship[6].

PCA transforms a multidimensional variable space to a space of principal components (PCs) that are linear combinations of the original variables (Jolliffe 1986). These new variables, the principal components, are linearly uncorrelated to each other and sequentially account for decreasing amounts of variance. By focusing on the first few components, the main sources of variation in the dataset and the influence of different objectives in discriminating between alternatives can be explored.

Using PCA for exploring MCDA problems has been proposed several times (e.g. Raveh 2000, Losa et al. 2001, Hodgkin et al. 2005, Miettinen 2014), and perhaps most notably in PROMETHEE, where the GAIA (Geometrical Analysis for Interactive Assistance) plane supports visualisation of results (Mareschal and Brans 1988). Similarly, PCA has the potential to support hierarchy building as it also indicates how a model can be made more concise without losing information. It has already been used in sustainability assessments, where a key challenge is to keep the set of indicators as small as possible while still meeting the needs of users (Shiau et al. 2015).

3.3.2 Interpretation of the results of statistical analyses

[5] The presented statistical analyses are not appropriate if value functions are ordinal or if the preference model is not additive (i.e. if it is a non-measurable/ non-cardinal value function).

[6] For ordinal values or if the relationship is suspected to be non-linear, rank order correlations can be calculated with Spearman's rho or Kendall's tau rank correlation coefficients (Kendall 1938).

The direction and strength of correlations indicate which trade-offs are potentially most or least critical for the outcome of the later MCDA. Critical trade-offs are those trade-offs where the allocation of weights between two objectives can have an impact on the ranking of alternatives. Critical trade-offs can be identified by analysing the ranking order and performance scores of the alternatives with regard to two objectives (or more quantitatively by doing correlation analysis and identifying the strongest negative correlations between two objectives). For example, objectives which have opposite ranking are those whose trade-offs are critical.

From the trade-off perspective, positively correlated objectives are less interesting because the later allocation of weights across these objectives does not significantly impact the MCDA outcome. A strong positive correlation means that alternatives perform similarly on both objectives (alternatives doing well/poorly on one objective are also doing well/poorly on another objective). However, it can indicate that (i) objectives are closely related or overlapping, suggesting that they could be merged (e.g. ecological quality of a lake and water quality of the lake); or (ii) the achievement of one of the objectives is a means to achieve the other objective i.e. if an alternative performs well/poorly on a means objective this will also lead to good/poor performance on the fundamental objective (e.g. phosphorus load of a lake and eutrophication of the lake); or (iii) that the impacts the objectives capture have a common cause (e.g. a poor sewer pipe rehabilitation can lead to groundwater contamination and also pipe failure).

A strong negative correlation between two objectives means that alternatives that perform well on one objective perform poorly on the other and vice versa. Strong negative correlations and high relevancy of the objectives (see 2.3.1) suggest that the trade-off between these objectives may be a key factor in the overall evaluation. It is essential that these critical trade-off questions are explicitly presented in the later MCDA process. They can be "hidden" if strongly negatively correlated lowest-level objectives are located in different hierarchy branches and hierarchical weighting is applied[7]. Correlation analysis and PCA can, therefore, reveal the potential need for, or benefit of restructuring the objectives hierarchy. Also, it can inform the decision analyst on how to design weight elicitation, e.g. by asking a consistency check question for strongly negatively correlated sub-objectives located in different hierarchy branches.

The relative number of positive and negative correlations gives a clue about the sensitivity of the results of an MCDA model. If the number of positive correlations is much higher than the negative correlations,

[7] In hierarchical weighting, weights are defined separately within each branch at each hierarchical level; the weight of the lowest-level objective is the product of its parental objectives and its weight relative to its siblings.

the result is less sensitive to changes in the objectives' weights and alternatives' scores than in the case where negative correlations dominate.

3.4 Sensitivity analyses

Local sensitivity analyses of objectives´ weights are commonly used in later MCDA stages to assess the impacts of uncertain preferences on the outcomes (Saltelli et al. 2004, Zheng et al. 2016). Payne et al. (1999) also suggest its use in all stages of the process. We used it firstly, to study how sensitivity analysis can support hierarchy building and, secondly, to analyse how sensitive the results are to the elimination of lowest-level objectives.

We performed two types of local sensitivity analyses:

i. Impact of lowest-level objectives' weights: we used the original hierarchy assuming equal weights for all lowest-level objectives because we aimed to investigate methods which could support hierarchy building before weights are assigned by stakeholders. For each objective in turn, the weight was reset to 0 and 0.25[8] and the associated overall ranks (i.e. at the top level of the objectives hierarchy) of the alternatives were determined and compared. The overall sensitivity was defined in terms of the changes in the rankings of decision alternatives (see procedures in S-2-6, S-2-7, S-3-3 and S-3-4).

ii. Impact of varying the number of objectives: the overall values of decision alternatives for three revised hierarchy options and all stakeholders were calculated and compared to their overall values using the original hierarchy. When eliminating objectives, the weights of the remaining objectives were renormalised across the hierarchy. We also conducted a complementary sensitivity analysis (or more precisely, uncertainty analysis) to investigate the effects of removing objectives when additionally all objectives' weights were considered uncertain (see S-3-8).

4 Case studies

In this section we illustrate, retrospectively, the potential of the range of methods presented in section 3 to inform the further development of objectives hierarchies in two case studies: sustainable water infrastructure planning in Switzerland (SWIP[9]) and regional land use planning in Finland (LUP). Two different cases were used to better understand the general applicability of the methods. These cases

---

[8] This is a plausible assumption of the possible weight range when no preference information is available. Our previous analysis covering 61 hierarchies (Marttunen et al. 2018) showed that the highest weight was in most cases between 0.1–0.25.

[9] http://www.eawag.ch/en/department/sww/projects/sustainable-water-infrastructure-planning-swip/

were chosen for three reasons: firstly,they used relatively large objectives hierarchies (19 and 16 lowest-level objectives); secondly, we were familiar with them as an author of this paper (different in each case) was a facilitator in the original study; and, thirdly, a substantial effort had been made when building the original hierarchies (for SWIP see Lienert et al. 2015). A noteworthy difference between the cases is that in the SWIP case many technical and natural attributes were used, whilst the LUP case used more constructed and proxy attributes.

The analyses were based mainly on the orginal data from the case studies. However, for the relevancy analysis additional data, which was not available from the original data set was required. We approached this by involving someone familiar with the topic (the SWIP case) or involved in the original study (the LUP case).

The Swiss case study, which was part of a project about Sustainable Water Infrastructure Planning (SWIP), focused on a wastewater infrastructure system decision. In the study area, 'Mönchaltorfer Aa' near Zürich, the three current wastewater treatment plants are running close to their capacity limits. Ten stakeholders were selected and participated in the MCDA process (Lienert et al. 2013, Zheng et al. 2016; 2018). Thirteen decision alternatives in four future scenarios were evaluated using objectives and attributes presented in Table 2. SWIP was carried out in 2010–2014 and the last author of this article was PI and responsible for the MCDA.

Table 2. Objectives, attributes and their ranges used in the Swiss Sustainable Water Infrastructure Planning case (SWIP, Lienert et al. 2015, Zheng et al. 2016, 2018).

Objective	Abbreviation	Description	Attribute	Worst alternative	Best alternative
1. Intergenerational equity					
Low future rehabilitation burden until 2050	Future_rehab	The term rehabilitation includes repairing damaged parts, their renovation, renewal, or replacement.	Realisation of the rehabilitation demand [%]	1 %	100 %
Flexible system adaption	Flexibility	The flexibility of a wastewater system refers to the difficulty of technical extension or deconstruction of the infrastructure.	Flexibility of technical extension or deconstruction of infrastructure [%]	20 %	85 %
2. Protection of water and other resources					
Good chemical state of water-courses	Chem_stat	Concentration of nutrients and pesticides does not exceed certain water quality limits at all reference points.	Average water quality across all reference points in catchment [0-1]	0.56	0.76
Low negative hydraulic impacts on surface water	Hydr_impact	Low discharge of stormwater from the infrastructure system is desirable, because large hydraulic impacts can destroy natural habitats (e.g. of fish).	Percentage of reference points in catchment that fulfil guidelines for stormwater handling [%]	7 %	16 %
Low contamination of ground water from sewers	Grw_nutr	Contaminants can reach ground water from leaking sewers (exfiltration of nutrients).	Water quality based on nutrients evaluated on [0,1] scale	0.75	0.78
Low contamination from infiltration structures	Grw_biocid	Biocides can reach ground water via infiltration structures (such as permeable pavement). Each biocide indicator was evaluated separately.	Water quality based on biocides evaluated on [0,1] scale	1	1
Recovery of nutrients	P_recovery	A good wastewater infrastructure makes efficient use of resources.	Recovery of phosphate from wastewater [%]	0 %	95 %

Objective	Abbreviation	Description	Attribute	Worst alternative	Best alternative
Efficient use of electrical energy	Energy	Electrical energy is needed mostly for treatment installations (aeration) and pumping. It is also possible to recover energy from the wastewater system.	Net energy consumption for wastewater treatment and transport [kWh/ person/ year]	310 kWh/ person/ year	5 kWh/ person/yea
3. Safe wastewater disposal					
Few gastro-intestinal infections through direct contact with wastewater	Illn_dir	If people get in contact with wastewater, they can get infected with pathogens. Direct contact is possible, if people that work with wastewater professionally or that maintain a decentralized treatment plant in their home do not follow precautionary measures.	Percentage of total population getting infected once per year[% / year]	0.11 %/ year	0.0005 %/ year
Few gastro-intestinal infections through indirect contact with wastewater	Illn_indir	The sewer system can be overloaded during heavy rain. Untreated wastewater can overflow into streets or houses, and direct contact is possible. People can get infected with pathogens if they swim in rivers or lakes after storms.	Number of CSOs / year /receiving water [no. / year / receiving water]	17	0
Few structural failures of drainage system	Drain_fail	If the sewers break down or are blocked because they are not well maintained, rain and wastewater can spill into e.g. streets or houses. This may be a nuisance, can disturb the traffic and business, or can damage property.	Weighted (by pipe diameter) number of pipe collapses and blockages per year and 1'000 inhabitants	8	0
Sufficient drainage capacity of drainage system	Drain_cap	During heavy rain untreated wastewater can overflow into streets or houses, and direct contact is possible.	Weighted (by city center and number of inhabitants) number of incidents of insufficient drainage capacity per year (e.g. overflowing of manholes)	1.049	0.074
4. High social acceptability					
Low unnecessary construction and road works	Road_disturb	Different infrastructures are currently sharing the underground. If the suppliers of these infrastructures collaborate when they are planning measures, there will be less unnecessary road and construction works.	Number of infrastructure sectors that collaborate in planning and construction	Lack of collaboration and many unnecessary construction sites and road works	Strong collaboratio and hardly any unnecessary constructior road works
High quality of management and operations	Mgmt_qual	The quality of management and operations of the wastewater system is evaluated by the EFQM Excellence Model, the most popular management quality tool in Europe. It assesses what an organization does, how it does it, and what it achieves.	Score of EFQM Excellence Model (European Foundation for Quality Management)	35 %	72 %
High co-determination of citizens in infrastructure decisions	Co_determ	The co-determination level, which means how much and in which way the citizens can take part in the decision making of infrastructure planning, also influences social acceptance.	Degree of co-determination [%]	10 %	90 %
Low time demand for end users	Time_dem	In conventional central wastewater treatment, professionals do all the maintenance and the end user does not have to invest any time. In the case of decentralized treatment units, end users have to invest time to operate and maintain the system, such as cleaning filters, or calling a service hotline.	Necessary time investment for operation and maintenance by end user [hours/ person/ year]	7 hours/ person/ year	0
Low additional area demand for end users	Area_dem	Decentralized wastewater 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 in the garden).	Additional area demand on private property per end user [m^2]	1.8	0
5. Low costs					
Low annual costs	Costs	Covers all expenses for the wastewater infrastructure system, including capital and running costs (operation and maintenance).	Annual cost per person in CHF	540 CHF/ person/ year	50 CHF/ per
Low cost increase	Cost_change	There are alternatives with low investment costs but high costs in operation and maintenance. On the other hand there exist alternatives with expensive start up investment costs, but resulting in little costs	Mean annual increase of costs in CHF / person / year until 2050	0.3 CHF/ person/ year	5.7 CHF/ pe

Objective	Abbreviation	Description	Attribute	Worst alternative	Best alternative
		for operation and maintenance.			

In the Finnish LUP case, the aim was to identify watersheds which would have the highest and lowest risk of water quality induced impacts if new peat extraction sites were established. The risk of negative impacts was considered highest when the water use values, water related nature protection values and the sensitivity of water bodies to additional loading were all high. The study covered 48 watersheds in Western Finland. The final hierarchy consisted of fourteen lowest-level objectives. Here we used an earlier version for demonstration purposes, including two extra objectives, flood risk and waterfowl hunting (Fig. 3).

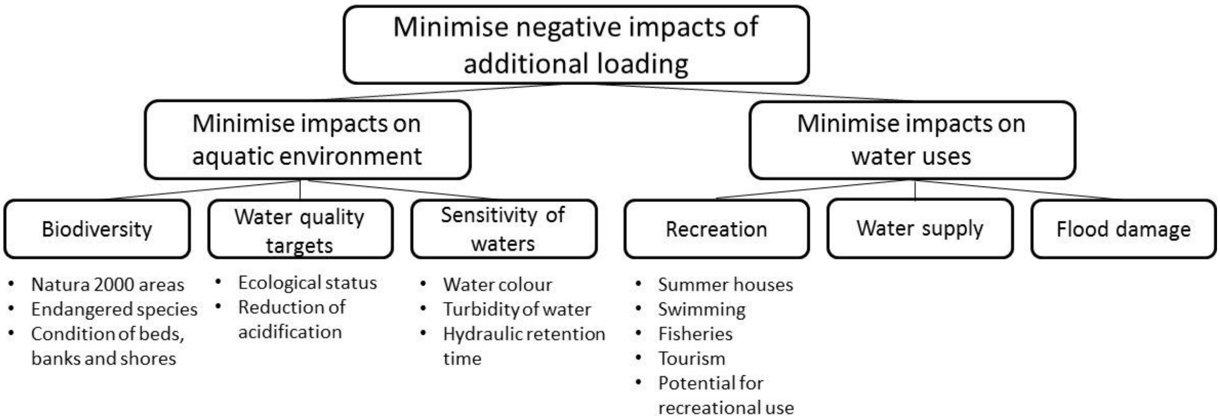

Figure 3. Objectives hierarchy in the Finnish land use planning case.

5. Illustration of the use of the methods in two retrospective cases

This section summarises the results of the *retrospective* simulation of the use of the methods discussed in Section 3 in these two cases. For simplicity, we present only the most important results from the SWIP case to illustrate the comprehensive analysis and to enable the reader to appreciate the value of the process and the main outcomes and contribution of each method. In section 5.2, we summarise how the results of the LUP case differed from the SWIP case. A more detailed description of the LUP case and results of the retrospective analyses are presented in the Supplementary material (S-2); the detailed results of the retrospective analyses of the SWIP case are presented in S-3.

5.1 Swiss wastewater infrastructure case (SWIP)

Means-end Analysis: in the SWIP case, we first constructed a means-ends network to illustrate the approach and to better understand the relationships between modelled objectives (ideally this analysis would be done before the development of an objectives hierarchy but it was not possible to do so in this retrospective illustration). This indicated that three of the objectives included in the hierarchy could be considered as means objectives: Quality of management and operations, Capacity of drainage system and Structural failures of drainage system (Tab. 2). All three objectives influence other fundamental objectives via several routes (Fig. 4). The classification of objectives into fundamental and means objectives is context dependent, presenting a starting point for further questions: (i) were certain objectives taken into account more than once, causing double-counting in the MCDA model? (ii) did these means objectives have further dimensions which were not covered by the fundamental objectives? and (iii) did these means objectives serve as operational and measurable proxies which were needed to assess the achievement of fundamental objectives? For reasons of space we do not further address these points, but they exemplify possible insights from such an analysis.

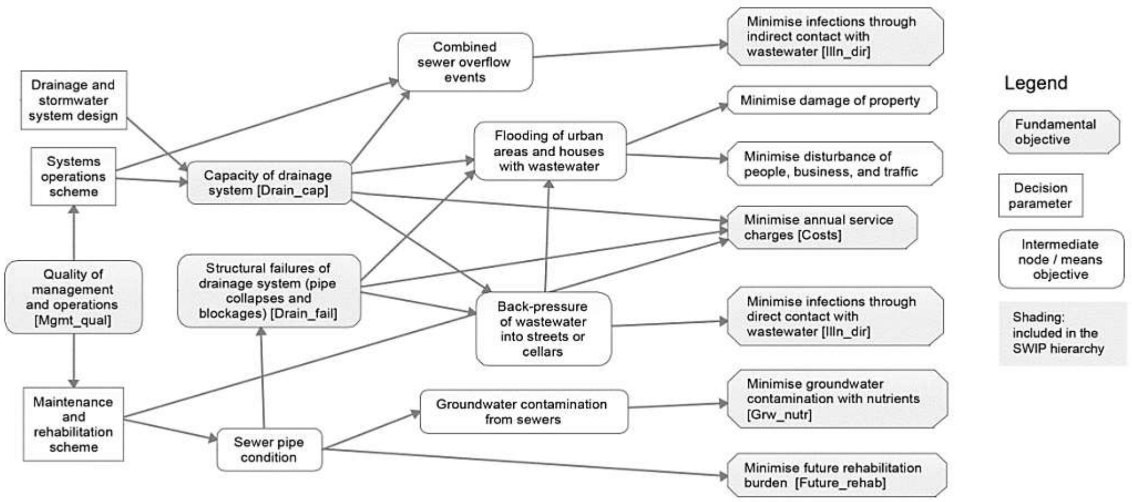

Figure 4. Part of a possible means-ends network for the SWIP case. For clarity, only those objectives which we initially considered as potential means objectives and the objectives closely related to them are presented.

Relevancy analysis: according to this analysis, five objectives belonged to the “Very high” relevancy class (Future_rehab, Illn_dir, Costs, Chem_stat, Drain_fail; top-right in Fig. 5). Two of them (Illn_dir, Costs) belonged to the “Very high” class both in terms of their importance and impact range. Three objectives were classified as having “Low” relevancy (Mgmt_quality, Road_disturbance, Cost_change). We emphasise, however, that this is a highly subjective analysis; a large diversity in the stakeholders’ weights in the original project suggest that different evaluators might have produced other outcomes. The visual categorisation of the objectives can inform discussions with stakeholders about the potential exclusion of objectives from the hierarchy. For instance, this visualisation might be especially useful if stakeholders consider some objectives to be of high general importance, while these objectives do not help to discriminate between decision alternatives in a specific application that is analysed with a subsequent MCDA. Such “high importance/ low impact” objectives could thus be removed in a next step (e.g. the groundwater objectives in the SWIP case: Grw_nutrient, Grw_biocide, Fig. 5).

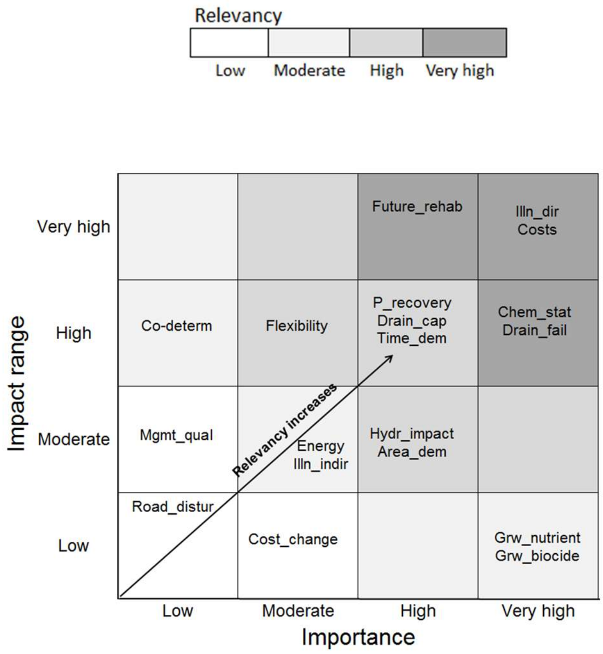

Figure 5. An illustrative example of the use of relevancy analysis in the Swiss wastewater infrastructure planning case (abbreviations see Tab. 2). [10]

Correlation and principal components analysis: there were 15 correlations exceeding 0.8 (total number of correlations in the SWIP case was 153, Fig. 6). The strongest positive correlations were between Area_dem, Drain_fail, Drain_cap and Illn_dir (abbreviations see Tab. 2) with a correlation coefficient of ≥0.9 for all pairwise combinations. The strongest negative correlations were between Energy and Illn_indir (-1.0) and between Costs and Drain_cap and Drain_fail and Grw_nutrient (-0.9, respectively). Additionally, four objectives (Flexibility, Energy, Co_determ and Costs) were negatively correlated with most others. Some objectives are strongly correlated because decision alternatives were generated from various measures using a strategy generation table (see Lienert et al. 2015) and this resulted in clusters of alternatives which have either a good or bad performance on several objectives. The number of objective pairs with relatively strong (>0.4) positive correlation (59) was three times higher than the corresponding number of negatively correlated objective pairs (20).

[10] The colours used in Figure 6 have been defined in the following manner:

1) Each importance and impact range class was assigned a numerical value so that Low = 0.25, Moderate = 0.50, High = 0.75, Very high =1.00.
2) For each cell, a relevance value was calculated by multiplying the corresponding values of importance and impact range (e.g. moderate importance & moderate impact range resulted in a relevancy value of 0.5*0.5=0.25).
3) The relevancy class of each cell was determined and colored according to the following threshold values: Low ≤ 0.15, Moderate > 0.15-0.3, High > 0.3-0.6, Very high > 0.6-1. The threshold values were set so that when the value of both axes are the same (e.g. moderate), the value of the cell in their crossing would also be the same (i.e. moderate relevancy).

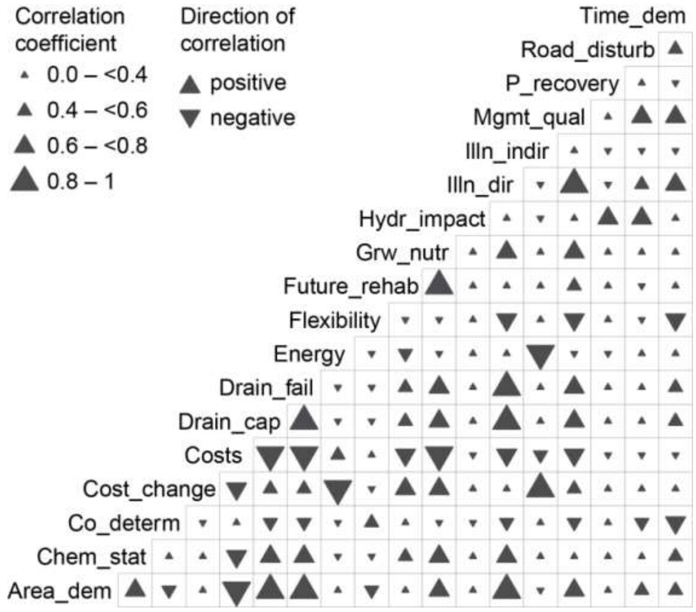

Figure 6. Results of the correlation analysis in the SWIP case (abbreviations see Table 2).

The PCA supports the insights from the correlation analysis (Fig. 7). The first component (PC 1) describes the main variation in the data set. The figure shows that Energy, Costs, Flexibility and Co_determ point in a different direction, highlighting their negative correlation with most other objectives. Based on this component alone, alternatives A4 and A5 can be clearly differentiated from the other alternatives. There are also three groups comprising several objectives pointing in the same direction (e.g. Road_disturb, Time_dem, Hydr._impact which point "northwest") suggesting that these are strongly positively correlated. Further analysis would be needed to evaluate whether they should all be included in the hierarchy.

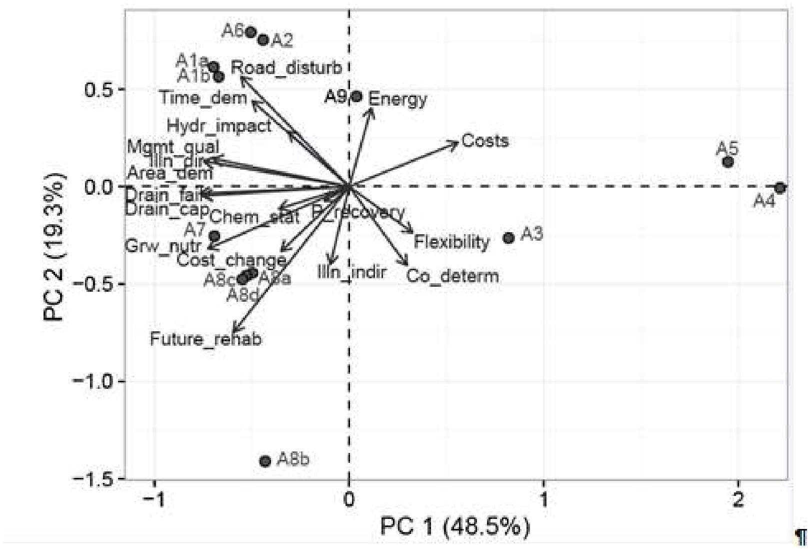

Figure 7. Results of the PCA for SWIP. A refers to decision alternatives; the other abbreviations refer to objectives (further explanations see Tab. 2).

Sensitivity analyses: we used local sensitivity analysis to investigate if there was a change to the set and/or the ranking of the five decision alternatives originally receiving the highest values when the weight of the objective ranged from 0 to 0.25. In general, the rankings of alternatives were insensitive to the weight changes. For twelve of the eighteen objectives there were no changes in rankings. The rankings were most sensitive to changes in the weights of Low time demand of the end-users, Low cost increase and Few gastro-intestinal infections through indirect contact with wastewater. For these objectives, the best decision alternative moved from the first to the fourth (Time_dem, Fig. 8) or from the first to the second rank (see S-3-3).

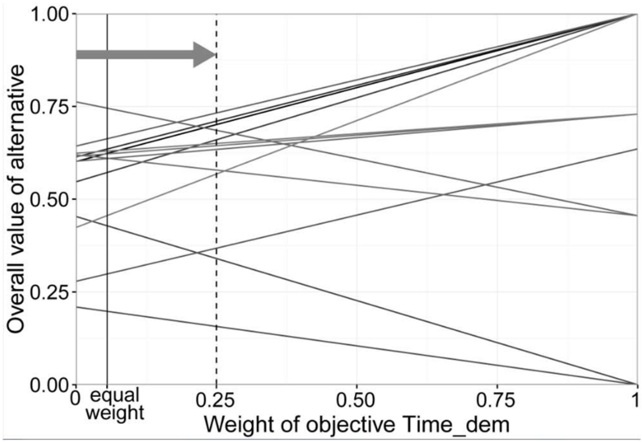

Figure 8. Example of local sensitivity analysis for SWIP – each line shows how the overall value of one of the 13 wastewater alternatives changes when the weight of the objective (here Low time demand for end users) was varied between 0 and 1.

Summary of findings: with regard to possibilities for simplifying the original hierarchy, both the means-ends network and statistical analyses suggest objectives which might be excluded. As the SWIP project was completed before this research was initiated, it was not feasible to follow up with the stakeholders who were originally involved to discuss the simplification of the hierarchy. Therefore, we used the results of the combined relevancy and sensitivity analyses to select three candidate sets of objectives which could be eliminated. We then analysed the sensitivity of the results (e.g. the ranking of the alternatives) to these changes to the objectives hierarchy.

In the development of the simplified objectives hierarchies, we eliminated those objectives having the lowest relevancy and sensitivity.

- REV-H1: Exclusion of the objectives High quality of management and operations (Mgmt_qual; low relevancy and low sensitivity), Low contamination of ground water from sewers (Grw_nutrient; low relevancy and low sensitivity) and Low unnecessary construction and road works (Road_disturb; low relevancy and moderate sensitivity).

- REV-H2: Exclusion of Mean annual increase of costs (Cost_change; low relevancy and moderate sensitivity), Low additional area demand for end-users (Area_dem; moderate relevancy, low sensitivity) and Efficient use of electrical energy (Energy; moderate relevancy and low sensitivity).
- REV-H3: Combines options 1 and 2 (six objectives were excluded).

We conducted sensitivity analyses regarding the outcomes of the MCDA for each hierarchy option and for all ten stakeholders (S-3-6 and S-3-7). The results showed that the ranking of the three best-performing alternatives was very stable. In the option where all six objectives were excluded, the ranking of the best alternative changed for only one stakeholder, for whom the first and second ranked options changed places. For seven stakeholders, there were no changes in the ranking of the three originally best-performing decision alternatives. For the other two stakeholders, 3rd and 4th alternatives changed places.

5.2 Finnish land use planning case (LUP)

There were no significant differences between the two cases in terms of our experience in using the methods. Therefore, the results of the LUP case are presented very briefly here. A more detailed description can be found in the Supplementary material (S-2).

The proportion of objectives belonging to the “Very high” relevancy class was higher in the LUP case than in the SWIP case; 11 of 16 objectives in LUP and 5 of 19 in SWIP. In LUP, the only objectives which received low relevancy ratings were Waterfowl hunting and Flood risk. It is noteworthy that these two objectives were excluded from the original hierarchy at an early stage of the analysis. Our relevancy analysis did not lead us to consider removing any additional objectives. It should be noted, however, that the relevancy analysis questionnaire was completed by the EIA consultant, and due to the subjective nature of this analysis other evaluators might have expressed different views.

The correlations between objectives were lower in LUP than in SWIP. In LUP there were no strong correlations (>0.8) and the proportion of correlations exceeding 0.4 was much lower than in SWIP. The correlation analysis showed that the number of positive correlations between objectives was twice the negative ones, suggesting that the results are rather insensitive to the elimination of the objectives. In the PCA, the decision alternatives were scattered indicating that there is a large variation in the characteristics of the 48 analysed watersheds.

The analysis did not reveal any trade-offs which indicate a need to pay particular attention in the weight elicitation and there were no strong negative correlations between the objectives located in different hierarchy branches.

6. Discussion

6.1 Methods and the MCDA process

The proposed methods provide a versatile set of analytical tools that can be applied, individually or complementarily, in different phases of MCDA (see Fig. 2) to support a systematic hierarchy development process and promote dialogue between different parties (Fig. 9). The methods are relevant in any situation where, in the initial phases, a large number of potential objectives or indicators are identified - particularly if performance information is readily available. Table 3 summarises the potential benefits and also the challenges associated with each method.

The methods help to identify the most relevant objectives and potential redundancies between the objectives. The systematic analyses can also improve the overall understanding of critical trade-offs (i.e. those trade-offs to which the ranking of alternatives is sensitive) and guide the decision analyst to explicitly focus on these. For instance, if we have two objectives which are strongly positively correlated, how the weight is distributed between them has much smaller impact on the outcome than if the objectives are negatively correlated.

The suitability and potential value of the methods depend on the nature and complexity of the decision situation and the quality of the available data. It is possible to apply only one method or several in a complementary way, as demonstrated in this study and illustrated in Figure 9. For instance, combining the results of the sensitivity analysis with those of relevancy analysis can give extra support for excluding an objective; if an objective has both low relevancy and low sensitivity, its exclusion will have lower impact on the MCDA results than if it has low relevancy and high sensitivity. Similarly, relevancy analysis could help to justify excluding objectives that are of high general importance to stakeholders but that do not discriminate between decision alternatives in MCDA ("high importance, low impact range"; see Fig. 5 for examples). This decision could be further informed if relevancy analysis is combined with PCA (Fig. 7). In practice, resource and time constraints may determine how much effort can be spent in hierarchy building. However, investing more resources in this phase can be worthwhile particularly in complex decision situations because a concise hierarchy can considerably reduce the effort for the later MCDA.

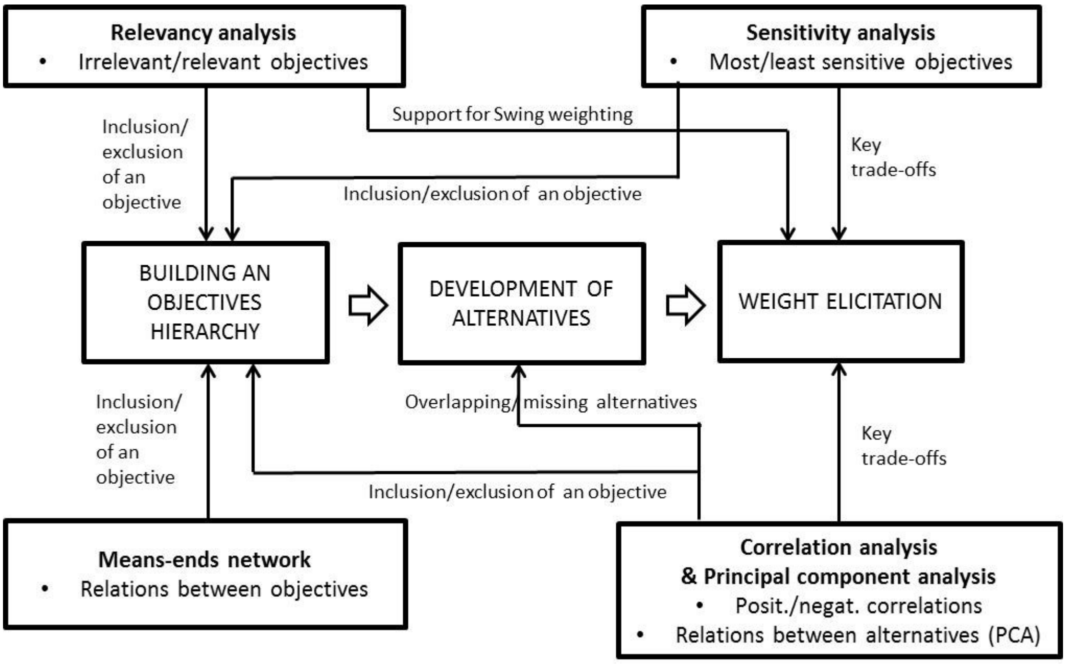

Figure 9. The proposed methods and how they can support different phases of MCDA.

There are also MCDA cases where alternatives are developed, refined and / or potentially screened out during the modelling process, which may also require, or enable, appropriate changes to the objectives hierarchy. For example, O'Brien and Brugha (2010) present a case where only two alternatives remained at the end of the analysis. These alternatives were similar on many of the objectives, which led to a 'magnifying glass' effect and a more-detailed focus on just one part of the objectives hierarchy.

Table 3. Potential benefits and challenges of the proposed methods.

Method	Benefits	Challenges
Means-ends network	Shows relationships among objectives (e.g. illustrates that different means can be used to achieve ends)	Clear understanding of the project's scope is required to develop an appropriate and useful network
	Supports interpretations of CA and PCA, as interactions in the system are identified	Differentiating means from ends can be challenging
Relevancy analysis	Provides a systematic framework for the screening of objectives	Designing and implementing the questionnaire can be laborious
	Encourages explicit consideration of the importance and impact range of objectives	People may have difficulties to fill in the questionnaire without support
	Helps to justify the evaluation and to communicate its results -	The wording of the questions can cause framing effect and anchoring bias
Correlation analysis (CA) / Principal component analysis	Helps to identify key trade-offs Results easy to understand (CA) Provides insights into the relationship between alternatives: allows detecting clusters of	Interpretation of results needs care, as statistical correlation does not necessarily imply a relationship between objectives Analysis generally requires the scores of alternatives

Method	Benefits	Challenges
(PCA)	alternatives that perform well on subsets of objectives or opposing alternatives with contrary performance on objectives (PCA)	Since the output of PCA needs careful interpretation, in most cases it will be a tool to inform the analyst rather than be useful for direct presentation to stakeholders Result can be sensitive to addition or removal of alternatives if ratio of number of objectives to alternatives is high
Local sensitivity analysis of weights	Routinely used and easy to perform Interpretation of the results is easy Dependencies between weights and alternatives' values are visible	Performing analysis objective by objective is laborious for large hierarchies Analysis generally requires scores of alternatives

6.2 Review and Reflection on the methods

6.2.1 Relevancy analysis

Relevancy analysis was developed to support the hierarchy building phase of MCDA when a large number objectives is initially identified and it is desirable to reduce their number for MCDA modeling. Using the terminology of Keeney (1992), relevancy analysis can be described as a value-focused thinking approach when it is applied in the early phases of MCDA without knowing the alternatives. On the other hand, if it is applied when alternatives and their impacts are known, it resembles the swing weight matrix method (Ewing et al. 2006, Parnell and Trainor 2009, Parnell et al. 2013) and could be described as an alternative-focused approach.

We used two completed environmental applications of MCDA as test beds, where we developed criteria, scales and questions to assess the importance, impact range and the relevancy of objectives. As our analysis was retrospective in nature, and it was possible for only one person in each case study (outside the research team) to respond to the questions, only tentative conclusions can be drawn.

In environmental applications, determining the relevancy of objectives has similarities with the impact significance assessment, which is a central phase of environmental impact assessment (Lawrence 2007, Wood 2008). Determining the significance of impacts requires that multiple impact dimensions are considered, such as intensity, duration, spatial distribution, likelihood and targeting of different groups of people (e.g. Lawrence 2007). The questions we used in relevancy analysis sought to make these different dimensions of importance and impact more visible and thus it encouraged their explicit consideration when evaluating the relevancy of the objectives. As specific knowledge may be needed to answer questions related to the general importance of objectives, experts from appropriate fields should be engaged.

In our cases, we asked similar questions for all impacts. These questions were most appropriate when assessing ecological, social or economic impacts, but in the Swiss SWIP project the respondent found this format to be less suitable for technical objectives. For instance, questions 1A-1C (Appendix 2) which relate to ecological, social and economic values of the objective, its sensitivity/ vulnerability and requirements of legislation were inappropriate in relation to "high quality of management and operations" and "flexible system adaptation". In the Finnish LUP project, assessing the difference in the alternatives' impact ranges (Question 2, Appendix 2) was challenging if proxy attributes were used (e.g., existence of significant flood risk areas as a proxy for flood damage. In future applications, it is recommended that the suitability of the questions is evaluated on an objective-by-objective basis and that they are modified or excluded where appropriate. An alternative approach would be to ask just two central questions: "What is the general importance of the objective?" and "How significant is the difference between the best and least good alternatives?" If only these two questions are used, it is important to present the impact information so that participants can take it into consideration when responding.

The potential gains of relevancy analysis are: (i) it provides a framework for a more transparent and understandable basis to identify the most relevant objectives; (ii) it can help to reveal areas of disagreement and thus support communication and participatory MCDA; and (iii) the visual illustration of the outcomes of relevancy analysis emphasises and can help stakeholders understand the two-dimensional nature of the weights of objectives (i.e. general importance and impact range). This has also been observed in the use of the swing weight matrix method (Ewing et al. 2009). Often stakeholders insist that objectives they consider important are included in MCDA even though the level of achievement of the objectives does not differ strongly between the alternatives. This can lead to overstated weights because consideration of the general importance of an objective may inappropriately dominate the impact range in weight elicitation. Relevancy analysis seeks to alleviate this problem by explicitly presenting the two key components of weight elicitation and thus allowing stakeholders to express which objectives they consider important in general, independently of which discriminate most between alternatives in the decision context.

We identified three potential challenges in the use of relevancy analysis. Firstly, a relatively large number of expert evaluations are needed to complete the analysis, i.e. five questions per objective in our cases. This requires extra effort; however, each expert is only asked to answer those questions within his/ her field of expertise. As the increased workload of relevancy analysis could be higher than the potential reduction achieved by reducing the hierarchy size, it may be useful to compare the extra effort needed for the relevancy analysis to its potential advantages before making the decision to use it. Secondly, the interpretation of general importance is different in different contexts and requires hard thinking (see also Parnell and Trainor 2009). Thirdly, though relevancy analysis provides an explicit differentiation of importance and impact, care should be taken to ensure that this is understood by participating stakeholders. A key distinction is that the importance assessment is inherently value-laden, but the impact range assessment should be based on facts. There are various ways of making the impact range assessment as transparent and value-neutral as possible. For instance, looking at the percentage change in the attribute (e.g. how much the best and worst alternatives deviate from the current state) or comparing the impact of the options in relation to some other human activities. For example, in the SWIP case, it could have been possible to compare the impacts of sewage to the impacts of nutrient load that comes from the agriculture in the case study area.

This study illustrates the potential of relevancy analysis. However, further research is needed to explore its implementation and effectiveness in multi-stakeholder settings. An interesting question is can relevancy analysis help to find a consensus regarding the objectives to be included in the hierarchy in cases where stakeholders' opinions vary substantially. Other areas for research include exploration and testing of the suitability of different types of questions for different types of objectives, and the potential to integrate relevancy analysis with the swing weight matrix method (Ewing et al. 2006, Parnell and Trainor 2009, Parnell et al. 2013) in the context of environmental decision-making and also in other application areas.

6.2.2 Correlation analysis and PCA

Correlation analysis and PCA can quickly provide an overview of the relationship between the alternatives' performances against different objectives in the MCDA model, supporting the identification of critical trade-offs or potential double counting. As previously mentioned, correlation analysis and PCA are primarily tools for analysts in their back-office work.

The added value of PCA over pairwise correlation is in providing an overview of the correlation structure within the whole dataset at once, not just between pairs of objectives. It does this in a way that best explains the variance in the data and enables the identification of the main positive and negative correlations at a glance through the PCA plot. Additionally, the alternatives are included in the plots which enables the analyst to visualise how/ if alternatives are differentiated by the principal components and which objectives contribute to the differentiation. Furthermore, it provides an indication of the effect that the removal of an objective may have on the discriminatory power of the objectives set.

In SWIP, the number of lowest-level objectives in the hierarchy branches varied from two to six (Fig. 6) which increases challenges related to hierarchical weighting and could potentially distort weights (see Marttunen et al. 2018). Questions to check the consistency of the weights across the branches of the hierarchy are recommended in the literature to diminish this risk (Hämäläinen and Alaja 2008, Morton and Fasolo 2009, Schuwirth et al. 2015). Strong negative correlations between alternatives' scores on different objectives indicate critical trade-offs. In future, we would focus questions to check the consistency of weights allocated to objectives associated with the strongest negative correlations across hierarchy branches, namely: Costs vs. Drain_cap, Costs vs. Area_dem, Costs vs. Grw_nutr and Future_rehab vs. Grw_nutr.

The results of these two cases indicate that the tested statistical methods can provide support for the insights gained from sensitivity analysis and means-ends networks. However, we found it difficult to draw conclusions regarding the inclusion or exclusion of objectives using only the information provided by the statistical analyses. Statistical analyses may have greater potential in the cases where there is very large amount of data available (big data) than in our cases where the number of alternatives and objectives was rather limited. As in any statistical analysis, it is important to confirm that the observed correlation is real and not an artefact of the method or data, which becomes more likely as the number of pairwise comparisons increases. Therefore, statistical analyses should be viewed as a starting point for more detailed deliberations. Graphical tools, such as means-ends networks, cognitive mapping and influence diagrams can be useful in better understanding the nature of the relationships between the objectives.

As indicated above, the statistical and sensitivity analyses are intended as an "off-line" activity, at least in the first instance, to help the analyst better understand aspects of the problem data structure. This enables them to provide effective support to the decision makers and ensure, for example, that critical trade-offs and sensitivities are highlighted and explored and "double-counting" of objectives does not inadvertently occur. It is not envisaged that these analyses would be shared in full with decision makers – only, if appropriate, those aspects which can be expected to inform their understanding of the problem.

6.2.3 Sensitivity analysis

The aim of sensitivity analysis is to highlight where outcomes of the analysis are particularly sensitive / insensitive to the weights allocated to specific objectives in order to inform discussion with the decision makers. Thus, sensitivity analysis can provide further information regarding the need to include an objective in the hierarchy (if relevancy and sensitivity are both low, then there is more support for exclusion of an objective than in the case where sensitivity is moderate or high).

We found traditional local sensitivity analyses, realised by separately changing each objective's weight, to be useful but also laborious when working with large hierarchies. Local sensitivity analysis has limitations since the effects of varying several parameters at the same time are not investigated. If it is of interest to explore the decision models more comprehensively, global sensitivity analysis should be used (e.g. Mustajoki et al. 2006). However, the high computational cost may limit its usefulness at early stages of the decision process (Scholten et al.2015). Both kinds of sensitivity analyses can guide the analyst in the process of determining which objectives have the greatest impact on the outcomes of the decision model and therefore deserve particular attention in the weight elicitation process. For instance, consistency check questions can be addressed to specifically explore the sensitivity of these objectives.

The evaluation of the different hierarchy options showed that the MCDA models used in the two cases were rather insensitive to the elimination of the objectives that we detected as candidates for exclusion. As the correlation analysis showed, in both cases there were many more positive correlations between objectives than negative ones. Thus, the relative insensitivity to the exclusion of objectives is logical.

7 Concluding remarks

The construction of an appropriate objectives hierarchy is key to effective MCDA and achieving a balance between completeness and conciseness is a challenging task, particularly in contexts which are data rich and there is a perceived need to ensure that "everything is taken into account". As a consequence, in complex decision situations, hierarchies can become very large, potentially leading to a situation where the essential objectives are intermingled with irrelevant ones and to overly laborious processes of value elicitation and subsequent analysis.

We have presented a set of qualitative and quantitative methods (means-ends networks, relevancy analysis, correlation analyses, sensitivity analysis) which can be used individually or complementarily to support the process of developing an appropriate objectives hierarchy in such a context. Each method has previously been demonstrated to be potentially beneficial in supporting the development and analysis of multi-attribute value models. Through the retrospective analysis of two completed case studies, we illustrate and evaluate how these methods can be used independently, or in an integrated way (Fig. 9) to inform and support analysts and facilitators in the development of an appropriately concise model in situations characterised by a broad range of perspectives and by extensive data.

We discuss opportunities in the use of methods in section 6.1 and present these visually in Fig. 9. We highlight the potential of these methods to reveal key aspects of structure in the data which defines the performance of options (for example, correlations which could reveal critical trade-offs or potential double-counting). This can inform the analyst and enable her to better support the decision makers in to help identify the most relevant objectives and to find potential ways to develop an appropriate objectives hierarchy.

We emphasise that hierarchy development process should be informed by knowledge of the potential advantages and disadvantages of simplifying an objectives hierarchy as discussed in section 2.5. It is not suggested that the methods presented should be used in a strict manner, neither should they replace nor diminish a facilitator's deeper thinking and aspiration to understand the decision context. This study should be seen as a step towards future exploratory and comparative research which can lead to a more systematic and better supported approach to structuring MCDA problems; a topic that has received too little attention to date. Testing with decision-makers and stakeholders involved in different types of MCDA applications is needed to better understand the full potential of and possible pitfalls associated with the use of the proposed methods.

Supplementary material

Additional information concerning the case studies, the analyses and the results, including a general questionnaire template for relevancy analysis can be found at XXX.

Acknowledgements

We are grateful to Pascal Bücheler for his valuable assistance in coding the MCDA model and conducting sensitivity analyses. We thank Anna-Mari Lehmonen and Philipp Beutler for answering relevancy analysis questionnaires. We are grateful for the excellent comments of three reviewers which have enabled us to make the paper more concise and understandable. The study was funded by the Eawag Directorate.

REFERENCES

Adelman, L., Sticha, P., J. & Donnell, M., L. 1986. An experimental investigation of the relative effectiveness of two techniques from structuring multiattributed hierarchies. Organizational Behavior and Human Decision Processes**,** 188-196

Bana e Costa, C. A. & Beinat, E. 2005. Model-structuring in public decision-aiding. Working paper LSEOR 05.79. The London School of Economics and Political Science. http://eprints.lse.ac.uk/22716/1/05079.pdf (accessed 3.11.2016).

Barcus, A. & Montibeller, G. 2008. Supporting the allocation of software development work in distributed teams with multi-criteria decision analysis. Omega-International Journal of Management Science, 36**,** 464-475. doi:10.1016/j.omega.2006.04.013.

Belton, V. & Stewart, T. J. 2002. Multiple criteria decision analysis - An integrated approach, Boston, Kluwer.

Belton, V. & Stewart, T. J. 2010. Problem structuring and Multiple Criteria Decision Analysis. In: Ehrgott, M., Figueira, J. R. & Greco, S. (eds.) Trends in Multiple Criteria Decision Analysis. Springer US.

Bojórquez-Tapia, L. A., Sanchez-Colon, S. & Florez, A. 2005. Building consensus in environmental impact assessment through multicriteria modeling and sensitivity analysis. Environ Manage, 36**,** 469-81. doi:10.1007/s00267-004-0127-5.

Bond, S. D., Carlson, K. A. & Keeney, R. L. 2008. Generating objectives: Can decision makers articulate what they want? Management Science, 54**,** 56-70. doi:10.1287/mnsc.1070.0754.

Bond, S. D., Carlson, K. A. & Keeney, R. L. 2010. Improving the generation of decision objectives. Decision Analysis, 7, 238-255. doi:10.1287/deca.1100.0172.

Borcherding, K. & Von Winterfeldt, D. 1988. The effect of varying value trees on multiattribute evaluations. Acta Psychologica, 68, 153-170.

Braunschweig, T., Janssen, W. & Rieder, P. 2001. Identifying criteria for public agricultural research decisions. Research Policy, 30, 725-734. doi:Doi 10.1016/S0048-7333(00)00122-0.

Brownlow, S. A. & Watson, S. R. 1987. Structuring multi-attribute value hierarchies. Journal of the Operational Research Society, 38, 309-317. doi:10.1057/jors.1987.52.

Brugha, C. M. 2004. Structure of multi-criteria decision-making. Journal of the Operational Research Society, 55, 1156-1168. doi:10.1057/palgrave.jors.2601777.

Brugha, C. M. 2015. Foundation of Nomology. European Journal of Operational Research, 240, 734-747. doi:10.1016/j.ejor.2014.07.042.

Buede, D. 1986. Structuring value attributes Interfaces, 16, 52-62. doi:10.1287/inte.16.2.52.

Clemen, R. & Reilly, T. 2014. Making hard decisions with decision tools, South-Western Publishin g Co.

Cohen, J. & Cohen, P. 1983. Applied multivariate regression/ correlation analysis for the behavioral Sciences, Hillsdale, NJ, Lawrence Eribaum Associates.

Eisenführ, F., Weber, M. & Langer, T. 2010. Rational decision making, Berlin, Springer.

Ewing, Jr, P. L., Tarantino, W., & Parnell, G. S. 2006. Use of decision analysis in the army base realignment and closure (BRAC) 2005 military value analysis. Decision Analysis, 3(1), 33-49. https://doi.org/10.1287/deca.1060.0062.

Franco, L. A. & Hämäläinen, R. P. 2016. Behavioural operational research: Returning to the roots of the OR profession. European Journal of Operational Research, 249, 791-795. doi:10.1016/j.ejor.2015.10.034.

French, S., Maule, J. & Papamichail, N. 2009. Decision behaviour, analysis and support, New York, NY, Cambridge University Press.

Goossens, A. J. M., Basten, R. J. I., Hummel, J. M. & Van Der Wegen, L. L. M. 2015. Structuring AHP-based maintenance policy selection. BETA publicatie: working papers. Eindhoven: Technische Universiteit.

Gregory, R. 2017. The troubling logic of inclusivity in environmental consultations. Science,Technology & Human Values 42(1), 144-165. doi:10.1177/0162243916664016.

Gregory, R., Failing, L., Harstone, M., Long, G., McDaniels, T. & Ohlson, D. 2012. Structured decision making: A practical guide to environmental management choices, Chichester, Wiley-Blackwell.

Gregory, R. & Wellman, K. 2001. Bringing stakeholder values into environmental policy choices: a community-based estuary case study. Ecological Economics, 39, 37-52. doi:Doi 10.1016/S0921-8009(01)00214-2.

Hämäläinen, R. P. & Alaja, S. 2008. The threat of weighting biases in environmental decision analysis. Ecological Economics, 68, 556-569. doi:10.1016/j.ecolecon.2008.05.025.

Hämäläinen, R. P., Luoma, J. & Saarinen, E. 2013. On the importance of behavioral operational research: The case of understanding and communicating about dynamic systems. European Journal of Operational Research, 228, 623-634. doi:10.1016/j.ejor.2013.02.001.

Hobbs, B. F. & Meier, P. 2000. Energy decisions and the environment: A guide to the use of multicriteria methods, Boston, Kluwer Academic Publishers.

Hodgkin, J., Belton, V. & Koulouri, A. 2005. Supporting the intelligent MCDA user: A case study in multi-person multi-criteria decision support. European Journal of Operational Research, 160, 172-189. doi:10.1016/j.ejor.2004.03.007.

Imperia. 2015. Guidelines for the systematic impact significance assessment. Imperia project report. University of Jyväskylä. https://jyx.jyu.fi/dspace/bitstream/handle/123456789/49498/Guidelines_for_impact_significance_assessment.pdf?sequence=1 [Accessed 27.12.2017]

Jolliffe, I. T. 1986. Principal component analysis, Berlin, Springer-Verlag.

Keeney, R. L., Renn, O. & von Winterfeldt, D. 1987. Structuring West Germany's energy objectives. Energy Policy, 352-362.

Keeney, R. L. 1988. Building models of values. European Journal of Operational Research, 37, 149-157. doi:10.1016/0377-2217(88)90324-4.

Keeney, R. L. 1992. Value-focused thinking a path to creative decisionmaking, Cambridge, Massachusetts, Harvard University Press.

Keeney, R. L. 2004. Framing public policy decisions. International Journal of Technology, Policy and Management, 4, 95-115.

Keeney, R. L. & Raiffa, H. 1993. Decisions with multiple objectives preferences and value tradeoffs, Cambridge University Press.

Keeney, R. L. and Gregory, R. 2005. Selecting Attributes to Measure the Achievement of Objectives. Operations Research 53, 1-11. doi.org/10.1287/opre.1040.0158

Kendall, M. G. 1938. A new measure of rank correlation. Biometrika, 30, 81-93. doi:Doi 10.2307/2332226.

Kleinmuntz, D., N. 1988. Decomposition and the control of errors in decision analytic models. Massachusetts Institute of Technology.

Langhans, S.,D., Resichert, P. & Schuwirth, N. 2014. The method matters: A guide for indicator aggregation in ecological assessments. Ecological Indicators: 494-507. http://dx.doi.org/10.1016/j.ecolind.2014.05.014

Langhans, S. D. & Lienert, J. 2016. Four common simplifications of Multi-Criteria Decision Analysis do not hold for river rehabilitation. Plos One, 11. doi:http://dx.doi.org/10.1371/journal.pone.0150695

León, O. G. 1999. Value-focused thinking versus alternative-focused thinking: Effects on generation of objectives. Organizational Behavior and Human Decision Processes, 80, 213-227. doi:DOI 10.1006/obhd.1999.2860.

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, 134-48. doi:10.1016/j.jenvman.2013.03.052.

Lienert, J., Scholten, L., Egger, C. & Maurer, M. 2015. Structured decision-making for sustainable water infrastructure planning and four future scenarios. EURO Journal on Decision Processes, 3, 107-140. doi:10.1007/s40070-014-0030-0.

Losa, F. B., Van Den Honert, R. & Joubert, A. 2001. The multivariate analysis biplot as tool for conflict analysis in MCDA. Journal of Multi-Criteria Decision Analysis, 10, 273-284. doi:10.1002/mcda.308.

Maida, M., Maier, K. & Obwegeser, N. Evaluation of techniques for structuring multi-criteria decision problem. CONF-IRM 2012 Proceedings, 2012.

Maier, K. & Stix, V. 2013. A semi-automated approach for structuring multi criteria decision problems. European Journal of Operational Research, 225, 487-496. doi:10.1016/j.ejor.2012.10.018.

Mareschal, B. & Brans, J. P. 1988. Geometrical representations for MCDA. European Journal of Operational Research, 34, 69-77. doi:Doi 10.1016/0377-2217(88)90456-0.

Marttunen, M., Lienert, J. & Belton, V. 2017. Structuring problems for Multi-Criteria Decision Analysis in practice: A literature review of method combinations. European Journal of Operational Research 263, 1-17. doi.org/10.1016/j.ejor.2017.04.041.

Marttunen, M., Belton, V. & Lienert, J. 2018. Are objectives hierarchy related biases observed in practice? A meta-analysis of environmental and energy applications of Multi-Criteria Decision Analysis. European Journal of Operational Research 1, 178-194. doi.org/10.1016/j.ejor.2017.02.038.

Marttunen, M., Weber, C., Åberg, U. & Lienert, J. 2019. Identifying relevant objectives in environmental management decisions: An application to a national monitoring program for river restoration. Ecological Indicators 101, 851-866. doi.org/10.1016/j.ecolind.2018.11.042.

Mathieson, G. 2001. Best practice for using assessment hierarchies in operational analysis – principles and practical experiences. OR Insight 14, 3-13. doi:10.1057/ori.2001.6

Miettinen, K. 2014. Survey of methods to visualize alternatives in multiple criteria decision making problems. Or Spectrum, 36, 3-37. doi:10.1007/s00291-012-0297-0.

Montibeller, G. & Belton, V. 2006. Causal maps and the evaluation of decision options—A review. Journal of the Operational Research Society, 57, 779-791. doi:10.1057/palgrave.jors.2602214.

Montibeller, G. & Von Winterfeldt, D. 2015. Cognitive and motivational biases in decision and risk analysis. Risk Analysis, 35, 1230-1251. doi:10.1111/risa.12360.

Montibeller Neto, G., Ackermann, F., Belton, V. & Ensslin, L. 2001. Reasoning maps for decision aid: A method to help integrated problem structuring and exploring of decision alternatives. ORP3 Paris, September 26-29. 2001.

Morton, A. & Fasolo, B. 2009. Behavioural decision theory for multi-criteria decision analysis: a guided tour. Journal of Operational Research Society 60, 268-275. doi:10.1057/palgrave.jors.2602550.

Mustajoki, J., Hämäläinen, 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. doi:10.1016/j.ejor.2005.02.070.

O'Brien, D. B, and Brugha, C. M. 2010. Adapting and refining in multi-criteria decision-making. Journal of the Operational Research Society. 61, 756-767.

Pöyhönen, M. & Hämäläinen, R. P. 1998. Notes on the weighting biases in value trees. Journal of Behavioral Decision Making, 11, 139-150. doi:10.1002/(SICI)1099-0771(199806).

Pöyhönen, M., Vrolijk, H. & Hämäläinen, R. P. 2001. Behavioral and procedural consequences of structural variation in value trees. European Journal of Operational Research, 134, 216-227. doi:10.1016/S0377-2217(00)00255-1.

Parnell, G.S. & Trainor, T.E. 2009. Using the swing weight matrix to weight multiple objectives. Proceedings of the INCOSE international symposium, Singapore, July 19-23, 2009.

Parnell, G. S., Bresnick, T. A., Tani, S. N & Johnson, E. R. 2013. Introduction to Decision Analysis. Handbook of Decision Analysis. DOI: 10.1002/9781118515853.ch1

Payne et al., (1999) Measuring constructed preferences: Towards a building code, J. Risk Uncertainty 19, 243-270.

Phillips, L. 1984. A theory of requisite decision models. Acta Psychologica 56, 29-48.

Rahman, M. A., Jaumann, L., Lerche, N., Renatus, F., Buchs, A. K., Gade, R., Geldermann, J. & Sauter, M. 2015. Selection of the best inland waterway structure: A multicriteria decision analysis approach. Water Resources Management, 29, 2733-2749. doi:10.1007/s11269-015-0967-1.

Raveh, A. 2000. Co-plot: A graphic display method for geometrical representations of MCDM. European Journal of Operational Research 125, 670-678.

Retief, R., Morrison-Saunders, A., Geneletti, D. & Pope, J. 2013. Exploring the psychology of trade-off decision-making in environmental impact assessment. Impact Assessment and Project Appraisal 31, 13-23. doi: 10.1080/14615517.2013.768007.

Rosenhead, J. & Mingers, J. 2001. Rational Analysis for a Problematic World Revisited: Problem Structuring Methods for Complexity, Uncertainty and Conflict, Chichester, London, John Wiley & Sons. Ltd.

Saltelli, A., Tarantola, S., Campolongo, F. & Ratto, M. 2004. Sensitivity analysis in practice: A guide to assessing scientific models, JohnWiley & Sons Publishers.

Selart, M. & Johansen, S. T. 2011. Understanding the role of value-focused thinking in idea management. Creativity and Innovation Management, 20**,** 196-206. doi:10.1111/j.1467-8691.2011.00602.x.

Shiau, T. A., Huang, M. W. & LIN, W. Y. 2015. Developing an indicator system for measuring Taiwan's transport sustainability. International Journal of Sustainable Transportation, 9**,** 81-92. doi:10.1080/15568318.2012.738775.

Scholten, L., Schuwirth, N., Reichert, P., Lienert, J. (2015) Tackling uncertainty in multi-criteria decision analysis – An application to water supply infrastructure planning. European Journal of Operational Research 242, 243-260. http://dx.doi.org/10.1016/j.ejor.2014.09.044.

Schuwirth, N., Reichert, P. & Lienert, J. 2012. Methodological aspects of multi-criteria decision analysis for policy support: A case study on pharmaceutical removal from hospital wastewater. European Journal of Operational Research, 220, 472-483. doi:10.1016/j.ejor.2012.01.055.

Slovic, P., B. Fischhoff and S. Lichtenstein (1986). The Psychometric Study of Risk Perception. Risk Evaluation and Management. V. T. Covello, J. Menkes and J. Mumpower. Boston, MA, Springer US**:** 3-24.

Srdjevic, Z., Bajcetic, R. & Srdjevic, B. 2012. Identifying the criteria set for multicriteria decision making based on SWOT/PESTLE analysis: A case study of reconstructing a water intake structure. Water Resources Management, 26**,** 3379-3393. doi:10.1007/s11269-012-0077-2.

Stewart, T. J. & Scott, L. 1995. A scenario-based framework for multicriteria decision-analysis in water-resources planning. Water Resources Research, 31**,** 2835-2843. doi:Doi 10.1029/95wr01901.

Terrapon-Pfaff, J., , Thomas Fink, T., Viebahn P., Jamea, E.,M. 2017. Determining significance in social impact assessments (SIA) by applying both technical and participatory approaches: Methodology development and application in a case study of the concentrated solar power plant NOORO I in Morocco. Environmental Impact Assessment Review 66, 138-150. https://doi.org/10.1016/j.eiar.2017.06.008.

Valle, R. & Climaco, J. 2015. A new tool to facilitate quantitative assessment of green activities - A trial application for Rio de Janeiro. Technological Forecasting and Social Change, 98**,** 336-344. doi:10.1016/j.techfore.2015.03.006.

Von Winterfeldt, D. 1980. Structuring decision problems for decision analysis Acta Psychologica, 45**,** 71-93. doi:10.1016/0001-6918(80)90022-0.

Von Winterfeldt, D. & Edwards, W. 2007. Defining a decision analytic structure. In: Edwards, W., Miles, R., F. & Von Winterfeldt, D. (eds.) Advances in Decision Analysis. Cambridge: Cambridge University Press.

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

Weber, M., Eisenführ, F. & Von Winterfeldt, D. 1988. The effects of splitting attributes on weights in multiattribute utility measurement. Management Science, 34**,** 431-445.

Wood, G. 2008. Thresholds and criteria for evaluating and communicating impact significance in environmental statements: 'See no evil, hear no evil, speak no evil'? Environmental Impact Assessment Review, 28**,** 22-38. doi:10.1016/j.eiar.2007.03.003.

Zheng, J., Egger, C. & Lienert, J. 2016. A scenario-based MCDA framework for wastewater infrastructure planning under uncertainty. Journal of Environmental Management, 183, 895-908 doi:10.1016/j.jenvman.2016.09.027.

Zheng, J. & Lienert, J. 2018. Stakeholder interviews with two MAVT preference elicitation philosophies in a Swiss water infrastructure decision: Aggregation using SWING-weighting and disaggregation using UTA^{GMS} European Journal of Operational Research 267, 273-287. https://doi.org/10.1016/j.ejor.2017.11.018.

Appendix 1. Evaluation criteria to assess the general importance of the objective in the case studies using relevancy analysis. See main text for more information.

Importance scale	Criteria
Low	At least one of the following criteria is met and none of the criteria is in the categories "Moderate" or higher. • There are no particular economic, social, cultural or nature values related to the objective under consideration in the target area. • The objective is not sensitive to changes (example: increase in traffic noise near an airport) or recovers quickly from human pressure. • There are no regulations (e.g. legislation) concerning the use or state of the objective. • Objective is not a key concern to any of the stakeholders.
Moderate	At least one of the following criteria is met and none of the criteria is in the category "High": • Economic, social, cultural or nature values of the target area/receptor are of moderate importance to the objective under consideration. • The receptor's sensitivity is moderate. • There are no binding regulations concerning the use or state of the objective. • Local people or other stakeholders are moderately worried about the changes in the state of the objective.
High	At least one of the following criteria is met: • Economic, social, cultural or nature values related to the objective under consideration are high. • The objective is sensitive to changes in the external environment and the recovery lasts long • There are binding regulations (e.g. legislation) concerning the use or state of the objective. • Local people or other stakeholders are worried about the changes in the state of the objective.
Very high	At least one of the following criteria is met: • Economic, social, cultural or nature values related to the objective under consideration are very high. • The objective is very sensitive to changes in the external environment and the recovery lasts very long or does not happen at all. • There are very strict regulations (e.g. legislation) concerning the use or state of the objective. • Local people or other stakeholders are very worried about the changes in the state of the objective. Conflicts are probable.
Unable to determine	There is not enough information to make the assessment.

Appendix 2. Example of the questions used in the relevancy analysis.

1. GENERAL IMPORTANCE OF THE OBJECTIVE **“X”** IN THE STUDY AREA

A. Are there specific nature, society or economy related values with respect to the objective and what is their importance? Values which are associated with the other objectives in the hierarchy should not be considered to avoid double-counting.

	Not relevant/ None	Low	Moderate	High	Unable to determine
Nature values	☐	☐	☐	☐	☐
Social values	☐	☐	☐	☐	☐
Economic values	☐	☐	☐	☐	☐
Comment:					

B. Are there international, national, regional or municipality level legislation or recommendations related to the objective and if so how strict are they?

	No	Non-binding regulations	Binding regulations	Unable to determine
International	☐	☐	☐	☐
National	☐	☐	☐	☐
Regional	☐	☐	☐	☐
Municipality	☐	☐	☐	☐
Comment:				

C. How easily is the objective affected or how well does it tolerate changes (e.g. are small changes significant, are there any particularly sensitive targets/objects in the study area)?

Low ☐ Moderate ☐ High ☐ Unable to determine ☐

Comment:

D. How important is it to improve the achievement of the objective or prevent the deterioration of its achievement in the study area from the current state? When giving your estimate take into account your answers to the questions 1A-1C.

i. Improvement of the achievement of the objective

Not important ☐ Slightly ☐ Moderately ☐ Important ☐ Very important ☐ Unable to determine ☐

ii. Prevention of the deterioration of the objective

Not important ☐ Slightly ☐ Moderately ☐ Important ☐ Very important ☐ Unable to determine ☐

Comment:

2. DIFFERENCE IN THE IMPACT RANGE

A. How large is the difference between the worst and best cases with regard to the objective? When giving your estimate take also into account if there are differences in the spatial extent or duration of the alternatives’ impacts.

Example:

i. Worst case: (E.g. the average value of the water quality indicator of all observation points in catchment is 0.56.)

ii. Best case: (E.g. the average value of the water quality indicator of all observation points in catchment is 0.76.)

Low ☐ Moderate ☐ High ☐ Very high ☐ Unable to determine ☐

Comment: