

Loughborough University Institutional Repository

Modelling multicriteria value interactions with Reasoning Maps

This item was submitted to Loughborough University's Institutional Repository by the/an author.

Citation: RODRIGUES, T.C. ... et al, 2017. Modelling multicriteria value interactions with Reasoning Maps. *European Journal of Operational Research*, 258 (3), pp.1054-1071

Additional Information:

- This paper was accepted for publication in the journal *European Journal of Operational Research* and the definitive published version is available at <http://dx.doi.org/10.1016/j.ejor.2016.09.047>

Metadata Record: <https://dspace.lboro.ac.uk/2134/23926>

Version: Closed access

Publisher: © Elsevier Ltd

Rights: This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Please cite the published version.

Modelling Multicriteria Value Interactions with Reasoning Maps

Teresa C. Rodrigues^{a}, Gilberto Montibeller^b, Mónica D. Oliveira^a, Carlos A. Bana e Costa^a*

^a CEG-IST, Centre for Management Studies of Instituto Superior Técnico, Universidade de Lisboa, Avenida Rovisco Pais, 1049-001 Lisbon, Portugal

^b Management Science and Operations Group, School of Business and Economics, Loughborough University, Leicestershire LE11 3TU, United Kingdom.

Authors' emails: T.C. Rodrigues (teresacrodrigues@tecnico.ulisboa.pt), G. Montibeller (G.Montibeller@lboro.ac.uk), M.D. Oliveira (monica.oliveira@tecnico.ulisboa.pt), C.A. Bana e Costa (carlosbana@tecnico.ulisboa.pt).

**Corresponding author*

Abstract: Idiographic causal maps are extensively employed in Operational Research to support problem structuring and complex decision making processes. They model means-end or causal discourses as a network of concepts connected by links denoting influence, thus enabling the representation of chains of arguments made by decision-makers. There have been proposals to employ such structures to support the structuring of multicriteria evaluation models, within an additive value measurement framework. However, a drawback of this multi-methodological modelling is the loss of richness of interactions along the means-end chains when evaluating options. This has led to the development of methods that make use of the structure of the map itself to evaluate options, such as the Reasoning Maps method, which employs ordinal scales and ordinal operators for such evaluation. However, despite their potential, Reasoning Maps cannot model explicitly value interactions nor perform a quantitative ranking of options, limiting their applicability and usefulness. In this article we propose extending the Reasoning Maps approach through a multilinear evaluation model structure, built with the MACBETH multicriteria method. The model explicitly captures the value interactions between concepts along the map and employs the MACBETH protocol of questioning to assess the strength of influence for each means-end link. The feasibility of the proposed approach to evaluate options and to deal with multicriteria interactions is tested in a real-world application to support the construction of a population health index.

Keywords: Ideographical causal maps; Multicriteria Analysis; Reasoning Maps; MACBETH; Value interactions.

1. Introduction

Causal maps with an ideographic nature, which represent subjective knowledge by depicting the beliefs and concerns of a decision-maker (DM) or groups of DMs (Eden & Ackermann, 1998b), have been extensively employed in Operational Research for supporting problem structuring and complex decision-making processes (Eden, 1994; Eden & Ackermann, 2001, 2013b; Franco & Lord, 2011). These ideographic causal maps (ICMs) represent means-end or causal discourses as a network of concepts connected by links denoting chains of arguments (Eden & Ackermann, 1998b). Such maps have several characteristics whose analysis provide important insights to DMs, for instance the presence of circular relationships, dynamic interactions and the existence of dilemmas caused by the presence of positive and negative links – for details see (Eden, 2004). It is thus not surprising that integrating ICM with multicriteria decision analysis (MCDA) has been considered as a new direction for research (Wallenius et al., 2008).

An important domain of application for ICMs is to support the development of multicriteria evaluation models (Bana e Costa et al., 1999; Belton et al., 1997; Franco & Lord, 2011; Montibeller & Belton, 2006), as these maps provide valuable information about the structure and content of the value tree employed to evaluate decision options within an additive value measurement framework. More specifically, modelling complex interactions among concepts by constructing means-ends chains of arguments enables the identification of value independent fundamental concepts. These concepts can then be considered as the evaluation criteria for the assessment of decision options, assuming a simple additive value model.

A first drawback of this multi-methodological modelling is the loss of richness of interactions along the means-end chains when evaluating decision options (Montibeller & Belton, 2006; Montibeller et al., 2008). In addition, although simple additive value models are easy to understand and construct – possibly supported by ICMs – being a good compromise between simplicity and decision quality (Belton & Stewart, 2002), they are not suitable when there are value interactions that need to be explicitly addressed (von Winterfeldt & Edwards, 1986), as it happens in several real world problems.

The limitation concerning the loss of rich interactions along means-end chains has led to the development of methods that employ the structure of the map itself to evaluate decision options – instead of making a transition to a value tree – which provides a seamless integration between problem structuring and the evaluation of the options (Belton & Stewart, 2010; Montibeller & Belton, 2006). While smaller and simpler than the ICMs employed for problem structuring (Eden, 2004; Eden & Ackermann, 2001; Eden & Ackermann, 2004), therefore losing partially their richness, these maps are larger and more complex than the standard multicriteria models employed for the evaluation of options.

One of these methods, Reasoning Maps (Montibeller & Belton, 2009; Montibeller et al., 2008; Montibeller et al., 2007), employs ordinal scales and ordinal operators for such evaluation,

controlling the cognitive complexity of the elicitation process (Larichev, 1992; Montibeller & von Winterfeldt, 2015). However, and despite its potential (Belton & Stewart, 2010), the Reasoning Maps approach cannot either perform a quantitative ranking of options (Ma & Kinderen, 2016) or explicitly model value interactions, both limiting its applicability and usefulness.

In this article we thus propose to extend the Reasoning Maps approach, enabling a multilinear evaluation model structure that is built with MACBETH – the Measuring Attractiveness by a Categorical Based Evaluation Technique (Bana e Costa et al., 2012b) – which explicitly captures value interactions between concepts and provides a quantitative ranking of options. MACBETH asks DMs for qualitative information, thus controlling the cognitive complexity of the elicitation of preferences, but provides the power of inference of quantitative preference modelling. The proposed extended Reasoning Maps approach makes use of an alternative structure to the additive value model framework so as to model value interactions between different evaluation aspects, with the multilinear value model proposed in (Keeney & Raiffa, 1976) being adapted for its use in a Reasoning Map.

The potential benefits of the extended Reasoning Maps approach that we propose here are threefold. Firstly, they may provide insights and a better understanding of the evaluation problem, in comparison to the construction of an ICM followed by the development of a multicriteria additive model. Secondly, they increase the power of evaluation of Reasoning Maps. Thirdly, they provide a user-friendly way of modelling value interactions in multicriteria models.

The article is structured as follows. Section 2 briefly reviews ICMs and approaches that seek to evaluate decision options using an ICM structure. Section 3 presents the extended Reasoning Maps method and examples to illustrate the feasibility of applying the method in practice are outlined in section 4. Section 5 reflects on the real-world application of the method, emphasising the advantages and limitations of the proposed technique and concludes the paper.

2. Literature review

Causal maps have been employed for several purposes, among them, as problem structuring tools for decision aiding (e.g. (Ackermann et al., 2014), (Eden & Ackermann, 2004) and (Franco, 2013)), as representations of subjects' mental models for describing and comparing their models (e.g. (Bier, 2001), (Soetanto et al., 2011), (Wood et al., 2012) and (Zaksek & Arvai, 2004)), or as tools to capture expert knowledge about a structure of a system for causal reasoning simulations (e.g. (Azadeh et al., 2014), (Froelich et al., 2012), (Lopez & Salmeron, 2014) and (Papageorgiou & Froelich, 2012)). In the first and second type of application, causal maps have an idiographic nature (Eden & Ackermann, 1998a) and are focused on exploring a new and unfamiliar domain, while in the third researchers assume typically a nomothetic positioning (Montibeller et al., 2001), trying to confirm widely accepted and generalised assumptions related to a specific system.

ICMs (Ackermann, 2012; Eden, 2004) can be grounded on different theoretical basis (Eden, 1992; Marchant, 1999), there being no universal guidelines to their analysis (Eden & Ackermann, 1992) since

it depends on the purpose of the research that is being conducted as well as on the theoretical basis that is being followed.

For problem structuring methods, ICMs are often built up with a small group of key stakeholders or DMs (Ackermann et al., 2014; Eden & Ackermann, 2013a; Franco, 2013). These maps are usually large, with hundreds of concepts and complex interrelationships between them which may include loops and other types of relations beyond perceived positive/negative causal or influence connections (Eden, 2004). ICMs are employed to promote a shared understanding of the problem as they support the negotiation of a group's view of their world (Eden & Ackermann, 2013a). Several types of analysis based on the content and topological characteristics of the map are possible (Eden, 2004; Eden & Ackermann, 1998b). Such analyses, combined with insights emerging from the process of building the map (Ackermann et al., 2014), provide valuable support for structuring and framing ill-defined problematic situations. These maps can be constructed individually (creating cognitive maps) and merged into a single representation (cause map), or else developed interactively with the group, creating on-the-spot a causal map (Ackermann, 2012).

ICMs have long been used to support the evaluation of decision options – see (Belton & Stewart, 2010; Franco & Montibeller, 2011; Montibeller & Belton, 2006) for a detailed discussion –, either by analysing the map's topological characteristics, in particular the potency of an option or its shortest path to the goals (Eden, 2004) or by combining it with MCDA (Belton et al., 1997; Belton & Stewart, 2010).

The most common way of combining MCDA with an ICM is to build an external multicriteria model by extracting concepts from a large causal map (Belton et al., 1997; Franco & Lord, 2011; Silva et al., 2013; Walshe & Burgman, 2010), sometimes 'collapsing' the network into its main concepts (Bana e Costa et al., 1999; Bana e Costa et al., 2014; Ensslin et al., 2000). This type of analysis provides information about the structure and content of the value tree employed to evaluate decision options within an additive value framework. This way of modelling is simple but in some contexts can be problematic, specifically under the presence of several interdependent aspects and when it is important to explicitly model multidimensional quantitative performances. Instances where these challenges occur can be found in healthcare, performance measurement, or environmental management – see, for example, (Rodrigues, 2014), (Clivillé et al., 2007), and (Pinar et al., 2014). In such cases, as DMs are not willing to discard many of the initial indicators, it is not possible to derive a set of evaluation criteria from the ICM and use a simple additive model. In addition, this type of multicriteria evaluation model loses the rich structure of an ICM.

To overcome these drawbacks, there have been proposals to combine both methods – ICMs and MCDA – employing the structure of the map itself to evaluate decision options (Montibeller & Belton, 2006). This prevents the challenging transition between the two methods (Belton et al., 1997; Belton & Stewart, 2010) and makes fewer modelling assumptions than a standard additive multicriteria model as such combination permits to address explicitly value interactions between concepts (Montibeller & Belton, 2006).

The use of an ICM structure to perform a multicriteria evaluation of decision options demands for enriched information about the strengths of links, which helps in extending the power of inference of ICMs. However, this typically requires much smaller maps with collapsed and simplified structures, which are usually derived from the large map employed to support problem structuring (Montibeller & Belton, 2006). These maps do not have the richness of a large ICM, as they are focused on the evaluation of alternatives.

Within this perspective, some methods have been proposed in the last decade, as can be observed in the summary of studies in Table 1.

Approach by	Dynamic/static analysis	Indetermination and/or indistinction	Strength of links	No options or options as nodes or options via descriptors
(Tzeng et al., 2010)	Dynamic	Addresses indetermination and indistinction	Likert-scale	No options
(Yu & Tzeng, 2006) (Xiao et al., 2012)	Dynamic	Addresses indetermination and indistinction	Strength of links is assumed	No options
(Azadeh et al., 2015) (Elomda et al., 2013) (Baykasoğlu & Gölcük, 2015)	Dynamic	Addresses indetermination and indistinction	Fuzzy membership function	No options
(Michnik, 2013)	Dynamic	Addresses indetermination and indistinction	Likert-scale	Options as nodes
(Wellman, 1994) (Wellman, 1990)	Static	Does not address indetermination and indistinction	Qualitative scale	Options as nodes
(Kosko, 1986)	Static	Does not address indistinction; addresses indetermination	Qualitative scale	Options as nodes
(Montibeller et al., 2008)	Static	Does not address indetermination and indistinction	Qualitative scale	Options via descriptors

Table 1 – Methods employing ICM to evaluate options

There is a group of methods that attempt to model value interactions among concepts, using a dynamic simulation analysis until a steady-state is found. They quantify the links based on verbal information elicited from a DM that represents the intensity of an influence link, which is typically converted into a numerical value with the application of a Likert-scale (e.g. Michnik, 2013) or with a fuzzy function (e.g. Elomda et al., 2013). Some of these methods do not explicitly discuss how alternatives are evaluated – such as (Yu & Tzeng, 2006), (Elomda et al., 2013), (Baykasoğlu & Gölcük, 2015), (Tzeng et al., 2010), (Azadeh et al., 2015) and (Xiao et al., 2012). The exception is (Michnik, 2013) who proposed a method in which options are represented in the map.

Despite the potential benefit of a dynamic analysis, there are several drawbacks in these methods from a multicriteria analysis perspective (Belton & Stewart, 2002; von Winterfeldt & Edwards, 1986). Firstly, they employ verbal scales to measure influence that are typically converted arbitrarily into a numerical score – such as in (Michnik, 2013) and (Tzeng et al., 2010). Secondly, many of these methods are inspired by AHP/ANP – such as in (Yu & Tzeng, 2006) and (Xiao et al., 2012) –, which is a controversial method in MCDA (Bana e Costa & Vansnick, 2008; Dyer, 1990). Thirdly, with the

exception of (Michnik, 2013), options are not explicitly modelled and evaluated within the map structure – such as in (Yu & Tzeng, 2006), (Elomda et al., 2013), (Baykasoğlu & Gölcük, 2015), (Tzeng et al., 2010) and (Xiao et al., 2012). However, (Michnik, 2013) does not use well-defined descriptors of performances, thus increasing the ambiguity of evaluating the performance of options (Keeney & Gregory, 2005) and constraining the evaluation only to a small set of options displayed in the map (Bana e Costa & Beinat, 2005).

Another group of methods that employ the structure of the map itself to evaluate options suggest a static evaluation of options, which is common in multicriteria evaluations (Montibeller et al., 2008). Those maps have a hierarchical structure without any loops (Montibeller et al., 2008). The original method proposed by (Kosko, 1986) suggested modelling the strength of influence with ordinal scales (qualitative labels) and using suitable operators for such a scale (Montibeller & Belton, 2009). (Wellman, 1994) suggested representing probabilistically the links and propagating the influence of means over ends. (Montibeller et al., 2008) proposed the Reasoning Map method, by employing ordinal scales and qualitative operators to evaluate options along the network, with descriptors associated to bottom-level concepts and with links denoting perceived influence.

In the static analysis there are two major challenges for propagating options' scores along the network: indetermination and indistinction (Montibeller & Belton, 2006). Indistinction occurs when the scale employed to measure the strength of links is not rich enough to enable the overall ranking of options. Indetermination happens when there are conflicting positive and negative chains of arguments from the means to the end nodes. These three methods using static analysis are all limited to a certain extent in the way they cope with these challenges. (Kosko, 1986), which converts all links into positive ones, avoids indetermination while suffering from indistinction due to a low granularity scale. Wellman, by using positive and negative links (Wellman, 1990; Wellman, 1994), is confronted with both indetermination and indistinction, which is compounded by the use of ambiguous qualitative labels for probability (Budescu et al., 1988). In addition, these two methods were not designed to perform multicriteria evaluations.

The third static method, Reasoning Maps, is instead designed to perform multicriteria evaluation of options: 'A more recent approach, stimulated both by practical experience and theoretical considerations with regard to the linking of the two distinct methodologies of cognitive/causal mapping and multi-attribute value analysis, has been the development of Reasoning Maps' (Belton & Stewart, 2010, pp. 215-216). However, the Reasoning Maps method is also confronted by indetermination and indistinction issues, which limits its modelling power (Ma & Kinderen, 2016). If these issues could be addressed, Reasoning Maps would have the capability of structuring and modelling quantitatively value interactions among evaluation concepts as well as of performing a quantitative rank of options, while only asking for qualitative information. Trying to address those two challenges is the purpose of this research.

The indetermination problem can be solved by adopting the procedure of (Kosko, 1986) which converts all negative links into positive ones by introducing a dis-concept (Konar & Chakraborty, 2005; Kosko, 1986) which is defined by assuming that the logical opposite is equivalent to the psychological opposite (this simplification will be addressed in more detail in section 3.3.).

The indistinction problem is more complex, as it involves converting qualitative into quantitative information with a sound theoretical approach. In the literature there are three main approaches to do so: AHP/ANP based methods (Saaty, 2004), fuzzy membership functions (Greco et al., 1999) and the MACBETH approach – first proposed in (Bana e Costa & Vansnick, 1994) and updated in (Bana e Costa et al., 2012b) and (Bana e Costa et al., 2016). The AHP/ANP underlying axiomatic (Smith & von Winterfeldt, 2004) is considered by several decision analysts as unsound, being a controversial method in the field of MCDA. The use of membership functions also raises theoretical problems as they typically have arbitrary shapes and do not model DMs preferences directly (Belton & Stewart, 2002; French, 1984). The MACBETH approach enables the conversion of qualitative preference information into quantitative value scores (Bana e Costa & Chagas, 2004), through a sound procedure that does not have the aforementioned problems (Angelis & Kanavos, 2016; Belton & Stewart, 2002; Montignac et al., 2009). Hence, exploring the use of Reasoning Maps with MACBETH will enable us to solve the problem of indistinction while maintaining the idea of the Reasoning Maps method of only asking DMs for qualitative information about strengths of influence.

The next section details the extensions to the Reasoning Maps approach that we propose. To distinguish our approach from the original Reasoning Maps method proposed by (Montibeller et al., 2008), we will adopt from now on the name ‘Reasoning-MACBETH map’ (or, simply, ‘ReasoningMAC’) and we will refer to the proposed extended version of the Reasoning Maps method as the ‘ReasoningMAC method’.

3. Reasoning-MACBETH maps

In this section, we explain the ReasoningMAC method by detailing its underlying assumptions and overarching activities as well as explaining how it resorts to a multilinear model structure to capture interactions among concepts. A step-by-step description of the ReasoningMAC modelling steps is also presented.

3.1. ReasoningMAC method for decision aiding

ICMs as well as the Reasoning Maps method belong to the constructivist paradigm in decision aiding (Bana e Costa & Pirlot, 1997; Roy, 1993) as they intend to support a group of DMs and stakeholders in addressing a problematic situation of shared concern within a complex environment. Moreover, as MACBETH follows the constructivist principles of process consultation (Bana e Costa et al., 2012b), Reasoning Maps and MACBETH can be mixed in a theoretically coherent manner within a single method – the ReasoningMAC method – avoiding the problem of paradigm incommensurability (Mingers, 1997).

Under this constructivist paradigm and in line with the concept of ‘requisiteness’ (Phillips, 1984; Phillips & Bana e Costa, 2007), the ReasoningMAC method is designed to be used within a requisite ICM, i.e. a network whose structure and content is considered rich enough for the problem addressed and additional refinements do not generate new insights to perform a multicriteria evaluation. In other words, we are balancing between practicality and richness of the model, with the aim of having an approach to be used as a decision-aiding method for multicriteria evaluation of options. ReasoningMAC can be applied with a single DM or else to a group of DMs employing decision conferencing (Phillips & Bana e Costa, 2007). This article is primarily focused on the technical aspects of the ReasoningMAC method, with the social aspects of group decision support not being addressed in detail.

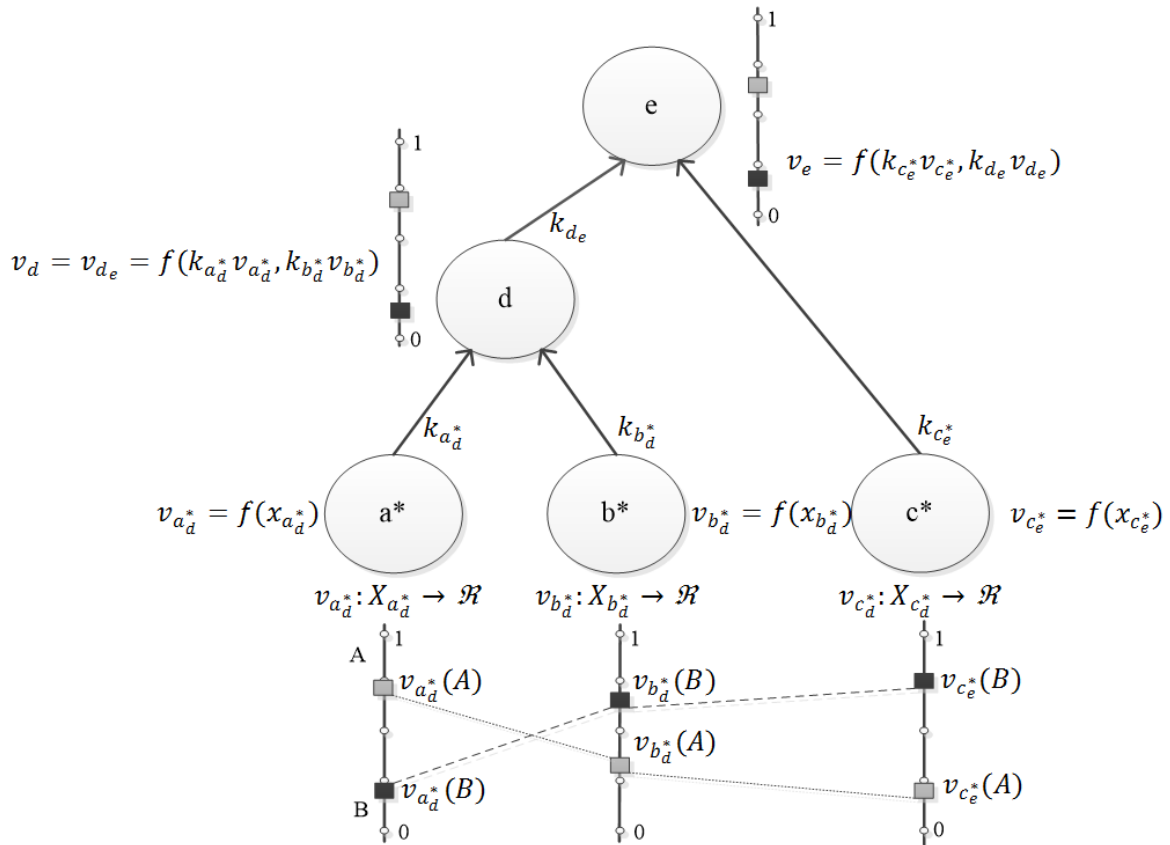


Fig. 1 – Illustrative example of the structure of a ReasoningMAC method evaluating options A and B. Note that, as all the links are positive, we omit the ‘+’ sign.

Based on the structure of the ICMs and with the purpose of performing a static analysis, a ReasoningMAC is a hierarchical network without negative links or circular relations, employed to evaluate options along the network. To clarify the structure of a ReasoningMAC, the following terminology will be adopted in the remaining of this article: a node that is influenced by another node is named *parent* node j and the node that influences a *parent* node j is named *child* node i_j ; a node that does not have in-arrows (being at the bottom of the network) is a special case of a *child* node being referred to as *descriptor* node (also called an *attribute* node). The *child* nodes that are also *descriptor* nodes will be signed with a ‘*’, i.e. i_j^* . A node which has no out-arrows is referred to as a *head* node.

Figure 1 depicts a ReasoningMAC to evaluate options A and B with five nodes: three *descriptor* nodes (a^* , b^* and c^*) and one *head* node (node e); the nodes represent concepts and are connected by arrows that represent influence; at the lowest level, each *descriptor* node i_j^* is assigned with a descriptor of performance $X_{i_j^*}$ (as defined below) and a value function $v_{i_j^*} = f(x_{i_j^*}) : X_{i_j^*} \rightarrow \mathcal{R}$, that converts the performance ($x_{i_j^*}$) of an option (A or B) into a real value score (respectively, $v_{i_j^*}(A)$ or $v_{i_j^*}(B)$) that measures the value of that option on the *descriptor* node. The value of an option (A or B) on a *parent* node j (respectively, $v_j(A)$ or $v_j(B)$), is a function of both the value of the option on each of the respective *child* nodes and the intensity of the influence of each arrow that connects each *child* node to the *parent* node j . That is to say that, given the performance and the value score of each option on each *descriptor* node (nodes a^* , b^* and c^*) and the strength of the influence link between each pair of nodes, the value scores of that option on the *descriptor* nodes can be ‘propagated’ to determine its value scores on the *parent* nodes. To do so, the underlying assumptions of a ReasoningMAC are:

- It is possible to operationalise each *descriptor* (*child*) node i_j^* by associating to it an ordered set $X_{i_j^*}$ of plausible performance levels – i.e., a ‘descriptor of performance’ (Bana e Costa et al., 1999) or an ‘attribute’ (Keeney & Gregory, 2005). The ordinal requirement imposed on $X_{i_j^*}$ ensures that it is ‘preferential independent’ of the remaining ones ($\bar{X}_{i_j^*}$) (Dyer & Sarin, 1979) and, therefore, it will be possible to build a value function $v_{i_j^*}(\cdot)$ for each *descriptor* node i_j^* that converts performance in added value for the DM.
- It is possible to assess the intensity/strength of the influence of each means-end link k_{i_j} ; a link between two nodes indicates that the *child* node i_j influences the *parent* node j , i.e., an increase of the value of an option on the *child* node generates an increase in the value of the option on the *parent* node; k_{i_j} reflects the magnitude of the influence of the *child* node i_j on the *parent* node j .
- The value of an option on the *parent* node j , $v_j(\cdot)$, is a function of both the value of the option on the respective *child* nodes and the intensity of influence from the links that connect the *child* nodes to the *parent* node j . Hence $v_j(A)$ depicts the impact that an option A has on the *parent* node j .

3.2. Modelling steps of the ReasoningMAC method

Following the Reasoning Maps method proposed by (Montibeller et al., 2008), the ReasoningMAC method has two main recursive phases: *structuring the map* and *evaluating options with MACBETH*. The *structuring* phase comprises the following steps: S1) building an ICM with the DM or group of DMs with the aim of capturing and representing, in a network of means and ends concepts, significant evaluation concepts that are important for the DM or group of DMs; S2) adapting the structure of an ICM to a collapsed and simplified ReasoningMAC that has no negative links; S3) constructing a descriptor of performance for each *descriptor* node of each sub-ReasoningMAC; and S4) performing tests to verify the condition of weak difference independence. The *evaluating with MACBETH* phase

consists of building a value function for each *descriptor* node (step S5) and determining the strength of each influence link (step S6). This latter phase uses MACBETH to measure the strength of the influence of each means-end link (step S6), by asking the DM to pairwise compare the influence of the *child* nodes of the same *parent* node, and finally to assess the value on each *parent* node (step S7). Overall, the ReasoningMAC approach is developed through a sequence of steps further detailed in the following sub-sections.

3.3. Structuring a ReasoningMAC

3.3.1. Building an ICM with the decision-maker

There is extensive and useful advice from the field of problem structuring methods on how to build ICMs, which will not be covered here – for details see (Eden & Ackermann, 1998b) and (Eden & Ackermann, 2001). They can be constructed individually and merged into a single ICM representation or developed interactively with a group, creating an ICM on-the-spot (Ackermann, 2012).

In the context of evaluating options within the map structure, ICMs should adopt a ‘requisite’ (Phillips, 1984; Phillips & Bana e Costa, 2007) structure, i.e. the map should have a collapsed and simplified structure with a content that is sufficient to represent the group (or individual) perspective and additional map refinements will not generate new insights for the purpose of performing a multicriteria evaluation. Once the map is built-up (Step S1) the analyst must confirm that it includes all fundamental aspects of the problem; this condition is guaranteed by the ICM constructing-process where it is ensured that the map is requisite and reflects the beliefs and concerns of the DMs (Eden & Ackermann, 1998b).

3.3.2. Transforming the ICM into a ReasoningMAC

After structuring the ICM it is necessary to adapt it into a ReasoningMAC (Step S2), with the aim of applying the multilinear model in a means-ends network (as will be described in Section 3.4.1). This procedure takes place according to the sequence shown in Figure 2:

- (a) Transform the negative arrows into positive ones (see Figure 2.a). According to several authors (Jetter & Kok, 2014; Konar, 2007; Konar & Chakraborty, 2005; Kosko, 1986), as each node of the network represents a concept, it is possible to define its logical negation, the so called ‘dis-concept’. Hence, by using the ‘dis-concept’ property and assuming that the logical opposite is equivalent to the psychological opposite, negative arrows can be transformed into positive ones – for details see (Kosko, 1986), (Konar & Chakraborty, 2005), (Jetter & Kok, 2014) and (Konar, 2007). This transformation should be done starting from the *head* nodes and then work down the network and perform the necessary changes as, from our perspective, an increase of *head* nodes should be avoided – see example in Figure 2.a.
- (b) Eliminate any redundant nodes (see Figure 2.b). This step avoids double counting and reduces the elicitation burden. It can be accomplished using the following procedure: when a *child* node *a* influences two different *parent* nodes (*b* and *c*) that are going to be *child* nodes of the same *parent*

node d , it is possible to determine directly the influence of the node a on the *parent* node d , as all the impacts of node a are propagated, through b and c , to the *parent* node d . Therefore, the *child* nodes of the *parent* node d are redundant for the evaluation and can be eliminated. This procedure can only be performed when the *parent* nodes (b and c) are influenced only by the same solo *child* node a .

- (c) Eliminate redundant links (Figure 2.c). This step removes direct links when indirect links are present from a *child* node to a *parent* node, following the procedure suggested by (Nadkarni & Shenoy, 2001, 2004). Thus if a node a affects a node c through b , then an arrow from a to c is redundant and increases the complexity of the representation and might be removed.
- (d) Break down the ReasoningMAC into sub-ReasoningMACs (see Figure 2.d). A sub-ReasoningMAC is a map that represents for each *parent* node j the *child* nodes i_j that influence the parent node. This is possible as the analysis is static and the map's structure is hierarchical without circular relationships. *Child* nodes can be replicated, as illustrated in the figure. However, notice that the definition of the descriptor depends on which *parent* node the *descriptor* node is influencing. For instance, the *descriptor* node b^* influences a *parent* node c and a *parent* node d . Suppose a company that wants to 'launch new products' (*child* node b^*) to 'increase profit' (*parent* node c) as well as to 'increase diversification and reduce market risk' (*parent* node d); the set of products that will have a *good* impact on increasing profit (for instance, products that take advantage of economies of scale and learning) are not necessarily the same set of products that will increase diversification and reduce market risk. Hence, if a *descriptor* node influences two different *parent* nodes the analyst must build a descriptor of performance for each *descriptor* node of each sub-ReasoningMAC (X_{b^*c} and X_{b^*d}).

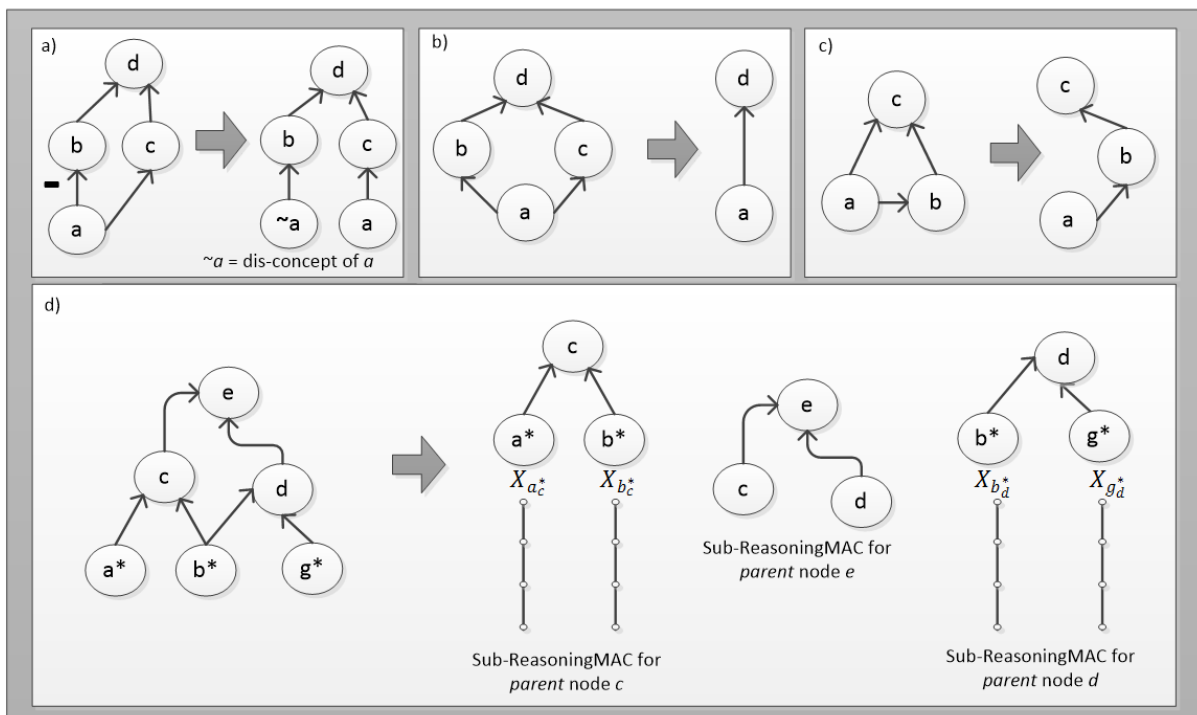


Fig. 2 – Procedure to adapt the ICM to a ReasoningMAC (all links are positive except the one with a minus sign).

3.3.3. Building a descriptor of performance for each descriptor node of each sub-ReasoningMAC

In building the descriptor of performance or attribute (Step S3), i.e. the ordered set of (quantitative or qualitative) plausible (observed or envisaged) performance levels that operationalise each *descriptor* (*child*) node, it is good practice to define the two end levels of performance, to bound the descriptor (von Winterfeldt & Edwards, 1986): a *best* level i.e., the performance on that *child* node necessary to imply the *best* performance on the respective *parent* node; and a *worst* level i.e., the performance on that *child* node necessary to imply the *worst* performance on the respective *parent* node. In addition, it is recommended (Bana e Costa et al., 2002) to identify also two levels of intrinsic value on the descriptor: a *good* level of performance i.e., the performance on that *child* node necessary to imply a *good* performance on the respective *parent* node; and a *neutral* level of performance i.e., the performance on that *child* node necessary to imply a *neutral* (neither attractive nor unattractive) performance on the respective *parent* node. Note that, in some cases some levels may overlap (e.g. when the *neutral* level is the same as the *worst* level).

3.3.4. Testing for the weak difference independence condition

The fact that, by construction, each descriptor $X_{i_j^*}$ is preferential independent of the remaining ones ($\bar{X}_{i_j^*}$), does not mean that $X_{i_j^*}$ is ‘difference independent’ of $\bar{X}_{i_j^*}$ – a necessary condition to construct a measurable additive value model (Dyer & Sarin, 1979). To test difference independence, firstly the analyst should fix two reference levels on each descriptor $X_{i_j^*}$, for instance, the *best* and *worst* levels or the *good* and *neutral* levels. We will illustrate the procedure for two *child* nodes m_h^* and p_h^* of a *parent* node h and the respective descriptors $X_{m_h^*}$ and $X_{p_h^*}$ in which: $G_{m_h^*}$ and $N_{m_h^*}$ are the *good* and *neutral* levels of m_h^* , respectively; and $G_{p_h^*}$ and $N_{p_h^*}$ are the *good* and *neutral* levels of p_h^* , respectively.

The idea behind the test is that, if the two *child* nodes m_h^* and p_h^* are difference independent then, on the *parent* node h : the difference in value between performance pairs $[G_{m_h^*}, G_{p_h^*}]$ and $[N_{m_h^*}, G_{p_h^*}]$ is equal to the difference in value between the pairs $[G_{m_h^*}, N_{p_h^*}]$ and $[N_{m_h^*}, N_{p_h^*}]$; and difference in value between performance pairs $[G_{m_h^*}, G_{p_h^*}]$ and $[G_{m_h^*}, N_{p_h^*}]$ is equal to the difference in value between the pairs $[N_{m_h^*}, G_{p_h^*}]$ and $[N_{m_h^*}, N_{p_h^*}]$. Strictly speaking, the additivity test, consists in verifying that:

$$v_h(G_{m_h^*}, G_{p_h^*}) - v_h(N_{m_h^*}, N_{p_h^*}) = [v_h(G_{m_h^*}, N_{p_h^*}) - v_h(N_{m_h^*}, N_{p_h^*})] + [v_h(N_{m_h^*}, G_{p_h^*}) - v_h(N_{m_h^*}, N_{p_h^*})] \quad (1)$$

Alternatively in practice (see section 4.2.2), the analyst can use MACBETH to test conditions (2) and (3) (the two are necessary because the difference independence relation is not symmetric) which are equivalent to condition (1):

$$[v_h(G_{m_h^*}, G_{p_h^*}) - v_h(N_{m_h^*}, G_{p_h^*})] / [v_h(G_{m_h^*}, N_{p_h^*}) - v_h(N_{m_h^*}, N_{p_h^*})] = 1 \quad (2)$$

$$[v_h(G_{m_h^*}, G_{p_h^*}) - v_h(G_{m_h^*}, N_{p_h^*})] / [v_h(N_{m_h^*}, G_{p_h^*}) - v_h(N_{m_h^*}, N_{p_h^*})] = 1 \quad (3)$$

Note that MACBETH can be used to test difference independence for more than two nodes (as exemplified in Appendix A of the Supplementary material).

In this paper we are interested in the situations in which difference independence does not hold and, consequently, the simple additive model cannot be applied. We assume that the analyst has opted for building a multilinear value model in the structure of the ReasoningMAC, starting by testing if the necessary condition of weak difference independence holds (Step S4).

Weak difference independence (Dyer & Sarin, 1979) between the descriptor $X_{i_j^*}$ of a node i_j^* and the remaining ones ($\bar{X}_{i_j^*}$) implies that the *relative* strengths of preference depend only on the performance levels of $X_{i_j^*}$ and not on the fixed performance levels on the remaining nodes. Consider a *parent* node j with n *descriptor (child)* nodes i_j^* . Consider also that the *descriptor (child)* node 1_j^* is *weak difference independent* of the other *child* nodes i_j^* for all $i_j^* \neq 1$ and, $x_{1_j^*}$ is a performance level of the descriptor of performance $X_{1_j^*}$ and $x_{i_j^*}$ and $x'_{i_j^*}$ are performance levels of the descriptor $X_{i_j^*}$. Then, it is possible to build a conditionally cardinal value function $v_{1_j^*}(\cdot)$ for the *descriptor (child)* node 1_j^* , by holding, at an arbitrary level, the performance on each one of the other *child* nodes. The value function $v_{1_j^*}(x_{1_j^*}, x_{2_j^*}, \dots, x_{n_j^*})$ is strategically equivalent to the value function $v_{1_j^*}(x_{1_j^*}, x'_{2_j^*}, \dots, x'_{n_j^*})$, i.e. the latter is a positive linear transformation of the former that depends only on $x_{i_j^*}$ for all $i_j^* \neq 1$ (Dyer & Sarin, 1979).

Weak difference independence can be verified with simple and intuitive tests – for details see (Currim & Sarin, 1984) and (Keeney, 1992). Consider the following test regarding *child* node i_j and three hypothetical options A, B and C: 1) holding the performance of these options on the remaining *child* nodes on the *neutral* level, the DM values the preference difference between options A and B as equal to the preference difference between B and C; 2) holding the performance of the remaining *child* nodes on other performance level, ask for the preference differences between options A and B and between B and C; 3) if the condition of weak difference independence holds, the answer in steps 1 and 2 must be the same, i.e., the preference difference between A and B is equal to the preference difference between B and C for all the values of the remaining nodes; otherwise this condition is not fulfilled.

In cases in which weak difference independence does not hold, the analyst can still opt for another type of non-additive model already applied to capture interactions with MACBETH as the 2-additive Choquet Integral (Brosig et al., 2016; Clivillé et al., 2007; Lopes et al., 2014). However this approach is out of the scope of this paper, and besides does not fit with the complex structure of a *ReasoningMAC* because it considers only interaction between pairs of nodes.

3.4. Evaluating options in a ReasoningMAC

3.4.1. Computing multilinear value along a ReasoningMAC

The application of the multilinear model along the ReasoningMAC allows the computation of a value score for a given option on a *parent* node (as illustrated in Figure 1). We adapt here the model proposed by (Keeney & Raiffa, 1976) because, firstly, it can be applied for a means-ends network and, secondly, it can be easily adapted to enable the explicit use of bipolar value scales that can be useful in complex problems (Grabisch & Labreuche, 2005a). In fact, for many practical cases, the value of an option on each concept is better expressed through a 'bipolar scale of value' (Rescher, 1969, p.64) in which the 0 corresponds to a *neutral* level of performance that enables to identify attractive options with a positive value score and unattractive options with a negative value score.

Therefore, we adapted the multilinear model (more details are available in Appendix B of the Supplementary material) as follows. Without loss of generality, using the notation previously defined, let $I_j = \{1_j^*, \dots, n_j^*\}$ be the set of *descriptor (child)* nodes of a *parent* node j , with i_j^* and s_j^* being two *descriptor (child)* nodes of the *parent* node j , $X_{i_j^*}$ the descriptor of performance of the *descriptor (child)* node i_j^* and $v_{i_j^*}(\cdot)$ the respective (conditionally cardinal) value function which is *conditional* to the performance levels of the other *child* nodes. Let $x_{i_j^*}$ denote the option's performance on the *descriptor (child)* node i_j^* , $v_{i_j^*}(x_{i_j^*})$ the respective value score and $x_{i_j^*}^+$ and $x_{i_j^*}^0$ the upper and lower reference levels, respectively; $[x_{1_j^*}, x_{2_j^*}, \dots, x_{n_j^*}]$ and $[v_{1_j^*}(x_{1_j^*}), v_{2_j^*}(x_{2_j^*}), \dots, v_{n_j^*}(x_{n_j^*})]$ are options performance and value profiles, respectively. Consider also that $k_{i_j^*}$ is the intensity/strength of the influence of *child* node i_j^* on *parent* node j , $k_{i_j^*s_j^*}$ is the synergic effect between nodes i_j^* and s_j^* and $k_{1_j^* \dots n_j^*}$ is the synergic effect between nodes $1_j^*, 2_j^* \dots$ and n_j^* . Then, the value of the *parent* node j is given by Eqs. 4-6:

$$v_j(x_{1_j^*}, \dots, x_{n_j^*}) =$$

$$\left\{ \begin{array}{l} \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) \text{ if } \left\{ \begin{array}{l} (\forall i_j^* \in I: v_{i_j^*}(x_{i_j^*}) \geq 0 \wedge \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) \geq \text{Max}\{Y_j\}) \vee \\ (\forall i_j^* \in I: v_{i_j^*}(x_{i_j^*}) \leq 0 \wedge \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) \leq \text{Min}\{Y_j\}) \vee \\ ((\exists i_j^*, s_j^* \in I: v_{i_j^*}(x_{i_j^*}) \cdot v_{s_j^*}(x_{s_j^*}) < 0 \wedge \text{Min}\{Y_j\} \leq \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) \leq \text{Max}\{Y_j\}) \end{array} \right. \quad (4) \\ \text{Max}\{Y_j\} \text{ if } \left\{ \begin{array}{l} (\forall i_j^* \in I: v_{i_j^*}(x_{i_j^*}) \geq 0 \wedge \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) < \text{Max}\{Y_j\}) \vee \\ ((\exists i_j^*, s_j^* \in I: v_{i_j^*}(x_{i_j^*}) \cdot v_{s_j^*}(x_{s_j^*}) < 0 \wedge \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) > \text{Max}\{Y_j\}) \end{array} \right. \quad (5) \\ \text{Min}\{Y_j\} \text{ if } \left\{ \begin{array}{l} (\forall i_j^* \in I: v_{i_j^*}(x_{i_j^*}) \leq 0 \wedge \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) > \text{Min}\{Y_j\}) \vee \\ ((\exists i_j^*, s_j^* \in I: v_{i_j^*}(x_{i_j^*}) \cdot v_{s_j^*}(x_{s_j^*}) < 0 \wedge \beta_j(x_{1_j^*}, \dots, x_{n_j^*}) < \text{Min}\{Y_j\}) \end{array} \right. \quad (6) \end{array} \right.$$

with,

$$\beta_j(x_{1_j^*}, \dots, x_{n_j^*}) = \sum_{i_j^*=1}^{n_j^*} k_{i_j^*} v_{i_j^*}(x_{i_j^*}) + \sum_{i_j^*=1}^{n_j^*} \sum_{s_j^*>i_j^*}^{n_j^*} (-1)^l k_{i_j^*s_j^*} \cdot |v_{i_j^*}(x_{i_j^*})| \cdot |v_{s_j^*}(x_{s_j^*})| + \dots +$$

$$+(-1)^l k_{1_j^* \dots n_j^*} \cdot |v_{1_j^*}(x_{1_j^*})| \cdot \dots \cdot |v_{n_j^*}(x_{n_j^*})|$$

$$Y_j = \left\{ v_j(x_{1_j^*}^0, x_{2_j^*}, \dots, x_{n_j^*}), v_j(x_{1_j^*}, x_{2_j^*}^0, \dots, x_{n_j^*}), \dots, v_j(x_{1_j^*}, x_{2_j^*}, \dots, x_{n_j^*}^0) \right\}$$

$$l = \begin{cases} 2, & \forall i_j^* \in I: v_{i_j^*}(x_{i_j^*}) \geq 0 \\ 1, & \text{otherwise} \end{cases}$$

$$i_j^*, s_j^* \in I, s_j^* > i_j^*$$

$$k_{i_j^*} > 0$$

where, $v_{i_j^*}(x_{i_j^*}^+) = 1$ and $v_{i_j^*}(x_{i_j^*}^0) = 0$ and $v_j(x_{1_j^*}^+, \dots, x_{n_j^*}^+) = 1$ and $v_j(x_{1_j^*}^0, \dots, x_{n_j^*}^0) = 0$.

Eqs. 4-6 can be interpreted (for details see Appendix B of the Supplementary material) as follows:

- For a *parent* node j whose *child* nodes have positive value scores: if $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ is lower than the best option of Y_j , than the value of the *parent* node j is given by the $Max\{Y_j\}$ (Eq. 5); otherwise, the value of a *parent* node j is modelled by $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ as defined in Eq. 4.
- For a *parent* node j whose *child* nodes have negative value scores: if $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ is higher than the worst option of Y_j , than the value of the *parent* node j is given by the $Min\{Y_j\}$ (Eq. 6); otherwise, the value of a *parent* node j is modelled by $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ as defined in Eq. 4.
- For a *parent* node j whose *child* nodes have positive and negative value scores: if $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ is higher than the best option of Y_j , than the value of the *parent* node j is given by the $Max\{Y_j\}$ (Eq. 5); if $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ is lower than the worst option of Y_j , than the value of the *parent* node j is given by the $Min\{Y_j\}$ (Eq. 6); otherwise, the value of *parent* node j is modelled by $\beta_j(x_{1_j^*}, \dots, x_{n_j^*})$ as defined in Eq. 4.

By applying Eqs. 4-6 bottom-up successively it is possible to compute the value score of an option on each node, as illustrated in Figure 1.

To adapt the multilinear model to be used for bipolar scales, we are assuming that, using the terminology of (Grabisch & Labreuche, 2005a), the ‘positive’ and ‘negative’ parts of the scale on each *parent* node are ‘symmetric’, so that the value on a *parent* node j of an option A with a value profile given by $[100_{1_j^*}, 0_{2_j^*}, \dots, 0_{n_j^*}]$ is the ‘opposite’ of the one of option B with the value profile given by $[-100_{1_j^*}, 0_{2_j^*}, \dots, 0_{n_j^*}]$; in other words $v_j(A) = -v_j(B)$ (for details see Appendix B of the Supplementary material). Notice that the extended multilinear model is generalised for bipolar scales but it can also be used for unipolar scales (in this case for all the *child* nodes of a given *parent* j node the extended multilinear model will be reduced to Eqs. 4 and 5).

3.4.2 Evaluating influences with MACBETH

The ReasoningMAC method enables the assessment of the value of every option on each *descriptor* node of each sub-ReasoningMAC, via a (conditionally cardinal) value function (Step S5) and, subsequently, by eliciting the strength of each influence link along the network (Step S6), the value of every option on each *parent* node (Step S7). For supporting such elicitations of preferences, we suggest using the MACBETH approach, which is assisted by the M-MACBETH decision support system (DSS) (Bana e Costa et al., 2015). One of the main advantages of MACBETH is that it only asks for qualitative pairwise comparison judgements of the difference of impact between stimuli. To facilitate the comparison between stimuli, the DM is asked to choose one of the seven qualitative categories (no difference, very weak, weak, moderate, strong, very strong and extreme). The selection of two or more categories is allowed. Each time a qualitative judgment is elicited the DSS verifies its consistency and offers suggestions to resolve inconsistencies if required. Regarding the construction of a conditionally cardinal value function, the M-MACBETH approach asks the DM to judge the differences of attractiveness between the various performance levels of the descriptor of performance by holding the value of the remaining *child* nodes at the lower reference level. The M-MACBETH derives a value scale from the consistent matrix of judgments that the DM must validate and adjust, if necessary – see examples of this process in (Bana e Costa & Oliveira, 2012), (Bana e Costa et al., 2002) and (Bana e Costa et al., 2012a). The MACBETH approach also supports assessing the intensity of each influence link in a ReasoningMAC: this is performed by pairwise comparing the influence of the *child* nodes of the same *parent* node, as will be illustrated in detail in section 4.1. A thorough explanation of the MACBETH approach and of its questioning protocols is provided by (Bana e Costa et al., 2012b).

4. Applying the ReasoningMAC method in practice

In this section we first illustrate the use of a ReasoningMAC to evaluate options, and subsequently detail how it was applied to help a DM group in modelling value interactions between evaluation aspects/concepts of a multicriteria model in a real-world problem.

4.1. An illustrative example

We will explain in this section how the ReasoningMAC method may help an individual to select his future car. We are assuming the need for modelling value interactions among concepts in this example. The application of the ReasoningMAC method begins with the construction of an ICM with the key aspects which the DM considers relevant for the choice of a *Good car* (step S1, leading to the ICM in Figure 3). The next step consists in adapting the ICM into a ReasoningMAC – step S2, as in Figure 3. Next, the ReasoningMAC must be divided into sub-ReasoningMACs, and a descriptor of performance (see Figure 3) is assigned to each descriptor node of each sub-ReasoningMACs (step S3). In this example, the performance levels *good* and *neutral* are defined as the upper and lower reference levels, respectively. The *good* and *neutral* levels will later be used to define hypothetical cars and by comparing their impact on each *parent* node, the parameters of the multilinear model will be

determined. The DM could decide to choose *best* and *worst* achievable levels to define these hypothetical options, but in most cases (Bana e Costa & Beinat, 2005; Pape, 2016) this can lead to unrealistic options, making the pairwise comparisons difficult or even unfeasible. Before starting the construction of the conditionally cardinal value scale for each *descriptor* node, the condition of weak difference independence must be verified with simple and intuitive tests (step S4). An example of these tests is presented in Table 2 – for details see (Currim & Sarin, 1984) and (Keeney, 1992).

After this structuring phase, we constructed a conditionally cardinal value function for each *descriptor* node of each sub-ReasoningMAC (step S5) employing the MACBETH approach. For example, for the node *Low price**, the DM is requested to qualitatively judge the perceived difference in impact (on the *parent* node *Affordability*) between pairs of performance levels of hypothetical cars, by holding the value of the node *High efficiency** in the *neutral* level (see Figure 4).

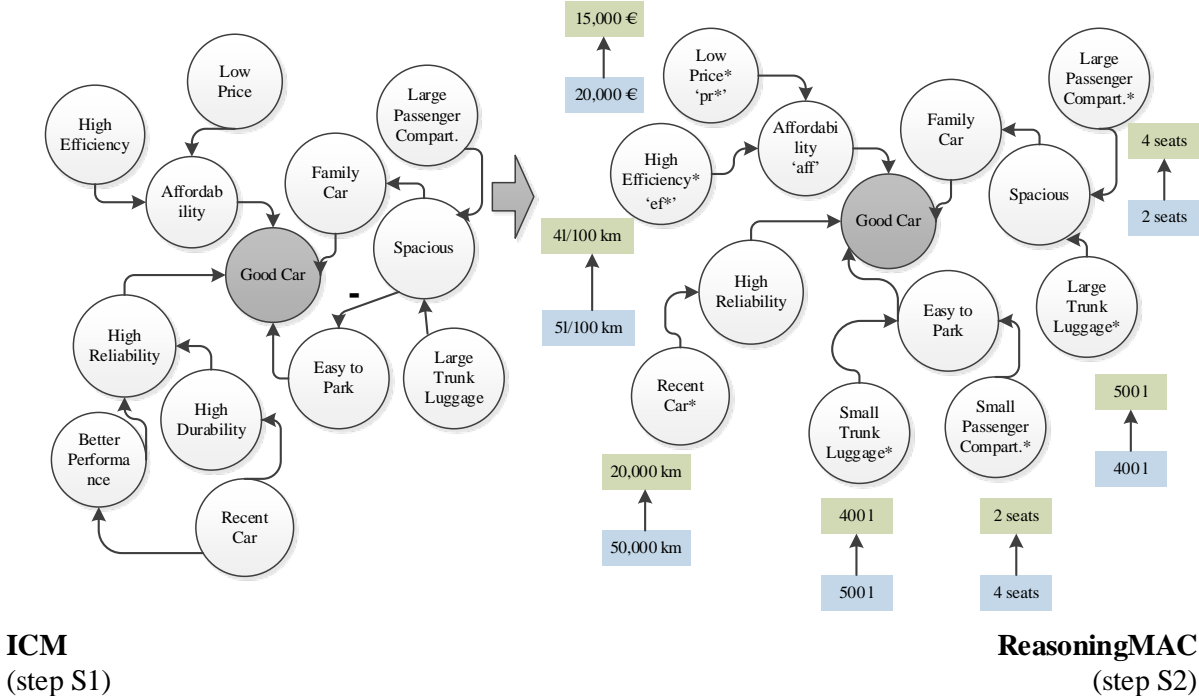


Fig. 3 – Converting an ICM into a ReasoningMAC. For each *descriptor* node the reference levels are specified.

Specify a cost *z* for car 2 such that your preference difference between car 1 and car 2 is the same as your preference difference between car 2 and car 3:

Car 1	Car 2	Car 3
5l/100 km	5l/100 km	5l/100 km
20 000 €	<i>z</i> ?	15 000 €

Similar provide a cost *t* for car 2:

Car 1	Car 2	Car 3
4l/100 km	4l/100 km	4l/100 km
20 000 €	<i>t</i> ?	15 000 €

If *z=t* and if it also holds true for all other values of price, then node *low price* is weak difference independent of the node *high efficiency*; note that weak difference independence is not a symmetrical relationship.

Table 2 – Testing *weak difference independence* between the node *low price* and the node *high efficiency* (notice that we employed the *good* and *neutral* levels to design the test, but other performance levels should be tested to ensure that this condition holds).

The next stage consists in eliciting the strength of the influence of each means-end link (step S6) again employing the MACBETH approach. As both *neutral* and *good* levels are already determined for each *descriptor* node of each sub-ReasoningMAC, one can construct hypothetical options for each *parent* node whose profiles are combinations of the reference levels (in this case the *neutral* and *good* levels) of the respective *child* nodes. For instance and following Figure 3, for the *parent* node *Affordability* (*aff*) the hypothetical options are defined by combining the reference levels of the *child* nodes *Low price** (*pr**) and *High efficiency** (*ef**) – i.e. by combining the *good* (*Gpr**) and *neutral* (*Npr**) performances of the node *Low price** and the *good* (*Gef**) and *neutral* (*Nef**) performances of the node *High efficiency** (*ef**) one can obtain four options [*Gpr**, *Gef**], [*Gpr**, *Nef**], [*Npr**, *Gef**], [*Npr**, *Nef**].

By pairwise comparing the impact of the options on the respective *parent* node, it is possible to determine the parameters of the multilinear model. For instance, regarding the *parent* node *Affordability*, the first question is phrased as follows: ‘how attractive/intense is the impact on *Affordability* of a car that costs 15,000 € and spends 5l/100 km – [*Gpr**, *Nef**] – comparing with the impact of a car that costs 20,000 € and spends 5l/100 km – [*Npr**, *Nef**]?’?. Consider that the DM responds with a *moderate* MACBETH qualitative judgement (see MACBETH matrix of Figure 5). A similar questioning protocol is subsequently used for each pair of hypothetical options, thus completing the MACBETH matrix in Figure 5. The M-MACBETH software then creates a value scale (displayed as a thermometer) with the scores of the hypothetical options compatible with the qualitative judgements, which can be adjusted and then validated by the DM. The same procedure described above must be performed to the *parent* nodes *Easy to park*, *High reliability* and *Spacious*.

Using these scores, the parameter(s) of the multilinear model can be assessed. From the analysis of the parameters of the multilinear model, one can understand the relationships between the components of the map as follows. Firstly, $k_{pr_{aff}^*} = 0.45$ is the influence of a hypothetical option that has a *good* performance on the node *Low price** – [*Gpr**, *Nef**] – thus capturing the intensity of the influence of the node *Low price** on the node *Affordability*. Note that if the reference levels of the node *Low price** change, then $k_{pr_{aff}^*}$ also changes, implying that the multilinear parameter would be different if the DM used other reference levels as the ‘upper’ and ‘lower’ references. Secondly, an hypothetical option that has a *good* performance on the node *Low price** – [*Gpr**, *Nef**] – has a higher impact on the node *Affordability* than an hypothetical option that has a *good* performance on the node *High efficiency** – [*Npr**, *Gef**]. Thirdly, the *child* nodes act conjunctively ($k_{pr_{aff}^*}e_{eff_{aff}^*} > 0$) (Grabisch & Roubens, 2000), i.e. the influence of a *child* node on the *parent* node is almost insignificant, but the influence of both *child* nodes is large.

Subsequently, the strength of the influence of the *child* nodes of the *parent* node *Good car* can be determined. It is unfeasible for a DM to consider four or more child nodes thus, following (Dyer & Lorber, 1982) and (Currim & Sarin, 1984), we suggest considering only two nodes at a time: firstly, the

influence of the node *Affordability* is compared with the influence of the node *High reliability* on the node *Good car*; then the joint-node *Affordability+High reliability* is compared with the node *Easy to park* (see Figure 6.a). This process is then repeated, until the influence of each *child* node on the *parent* node *Good car* has been determined, as depicted in Figure 6.a.

Overall, for a *parent* node *j* with *n* *child* nodes to determine the parameters with the multilinear model the DM can make a number of pairwise comparisons ranging from a maximum of $2^n(2^n - 1)/2$ to a minimum of $2^n - 1$; following the guidelines of (Bana e Costa & Chagas, 2004) it is recommended to input at least the border of the upper triangular portion of the MACBETH matrix (see Figure 5), in a total of $3(2^n - 2)$ pairwise comparisons judgements.

Low price* → Affordability	Level (thousand euros)
Best achievable	10
Good	15
Neutral	20
Worst achievable	25

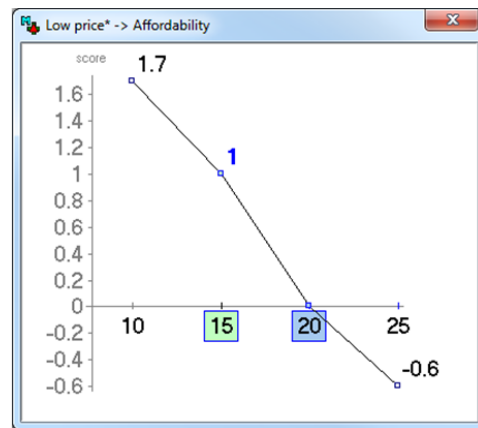
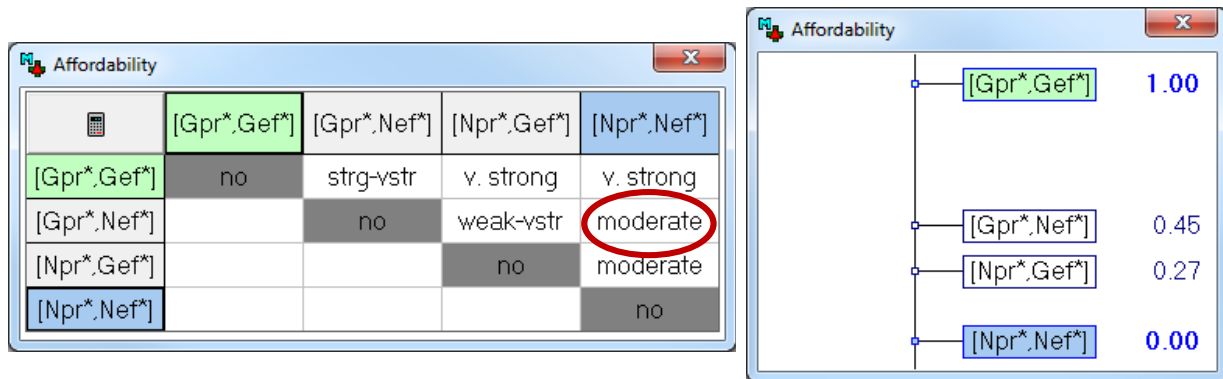


Fig. 4– Descriptor of performance the node *low price* and respective conditionally cardinal value function.



$$\beta_{aff}(x_{pr_{aff}^*}, x_{ef_{aff}^*}) = 0.45 v_{pr_{aff}^*}(x_{pr_{aff}^*}) + 0.27 v_{ef_{aff}^*}(x_{ef_{aff}^*}) + (-1)^l(0.28) |v_{pr_{aff}^*}(x_{pr_{aff}^*})| \cdot |v_{ef_{aff}^*}(x_{ef_{aff}^*})|$$

Fig. 5 – MACBETH matrix and respective thermometer for the *parent* node *affordability*; the scores of the hypothetical options were used to obtain $\beta_{aff}(x_{pr_{aff}^*}, x_{ef_{aff}^*})$ that is going to be used in Eqs. 4-6 to compute the value of the *parent* node *affordability*.

	Option A	Option B	Option C
Price	10 000 €	18 000 €	25 000 €
Efficiency	6l/100 km	4l/100 km	5l/100 km
Mileage	70,000 km	50,000 km	0 km
Number of seats	2 seats	4 seats	4 seats
Luggage compartment	420 l	350 l	500 l

Table 3 – Table of performances.

After determining the strength of the influence of each *child* node on the respective *parent* node, the multilinear model (Eqs. 4-6) can be applied in a recursively way to determine the impact of each option

on each *parent* node (step S7). For instance, consider that the DM wants to select his car out of three options (see Table 3). The application of the proposed model provides a rank of the options according to their impact on each *parent* node, particularly, in the *parent* node *Good car*, as in Figure 6.b. The use of the *neutral* and *good* references in each descriptor of performances positions the cars against these categories of intrinsic attractiveness.

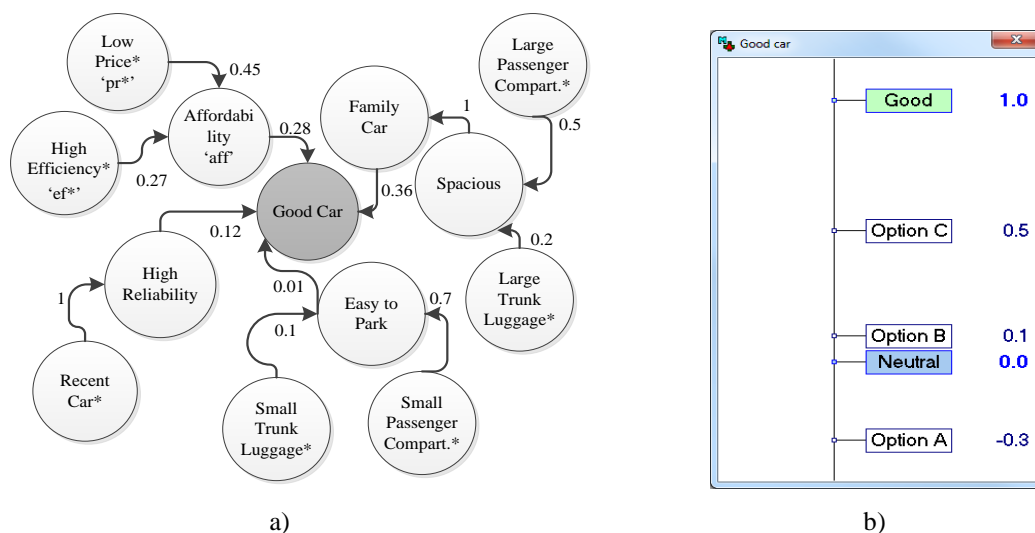


Fig. 6 – a) A Reasoning-MACBETH map with influences values. Note that synergic values are not specified in the map; b) thermometer of the options accordingly to their impact in the *parent* node *Good car*.

4.2. A real-world application of the ReasoningMAC method

We selected a case-study – in which it was paramount to model value interactions among several evaluation aspects – to test the feasibility of applying the method in practice. We thus applied the ReasoningMAC method into a real-world evaluation problem, where it was necessary to aggregate several performance indicators in an overall index to evaluate the health of populations of Portuguese municipalities (Rodrigues, 2014)¹. After providing some background information on the evaluation context we describe the process and take some insights from the application of the ReasoningMAC method.

4.2.1. Project background

The GeoHealthS project aimed at evaluating population health through the development of a population health index that was to be taken as the basis to monitor and assess population health on multiple dimensions and on aggregate at the Portuguese municipality level, and then to assess health inequalities along the Portuguese territory and the need to undertake health improvement policies. In this context, a social-technical process was designed to build a population health index: it integrated the technical elements of a multicriteria model, and the social elements of participatory methods, so that experts' views with different backgrounds could be considered.

The first step to build the additive model underlying the population health index consisted in structuring the set of (difference independent) evaluation criteria and corresponding performance

¹ Note that the case-study was developed by the first, third and fourth authors of this article.

descriptors to be used in a simple additive model. To define the set of evaluation criteria, the group of DMs started by discussing the contextual issues related to population health and specifying a set of health indicators that were relevant to appraise population health. From this discussion, it was then possible to define a value tree (Belton et al., 1997) with evaluation aspects organized within several areas of concern – a partial view of the tree is displayed in Figure 7. Each evaluation aspect is measured by one performance indicator. For some branches a simple additive model could be used (and evaluation aspects could be considered as an evaluation criterion in the additive model). However, for other branches, given the possible presence of value interactions, a simple additive model could not be employed, as it lacked theoretical significance (French, 1986) from a multicriteria value measurement perspective. In this latter case the group was not willing to discard any indicator from the initial set (as they are the ones that better reflect the phenomena under appraisal), even if potential value interactions between several indicators have emerged during the discussion. Hence, for these instances, a means-ends network was structured, and additivity between aspects was tested with the ReasoningMAC method – Figure 7 shows three (out of five) constructed networks where the ReasoningMAC method was used within a decision-conferencing process (Franco & Montibeller, 2010; Phillips & Bana e Costa, 2007). We select the ReasoningMAC II from Figure 7 to explain how the ReasoningMAC method was used.

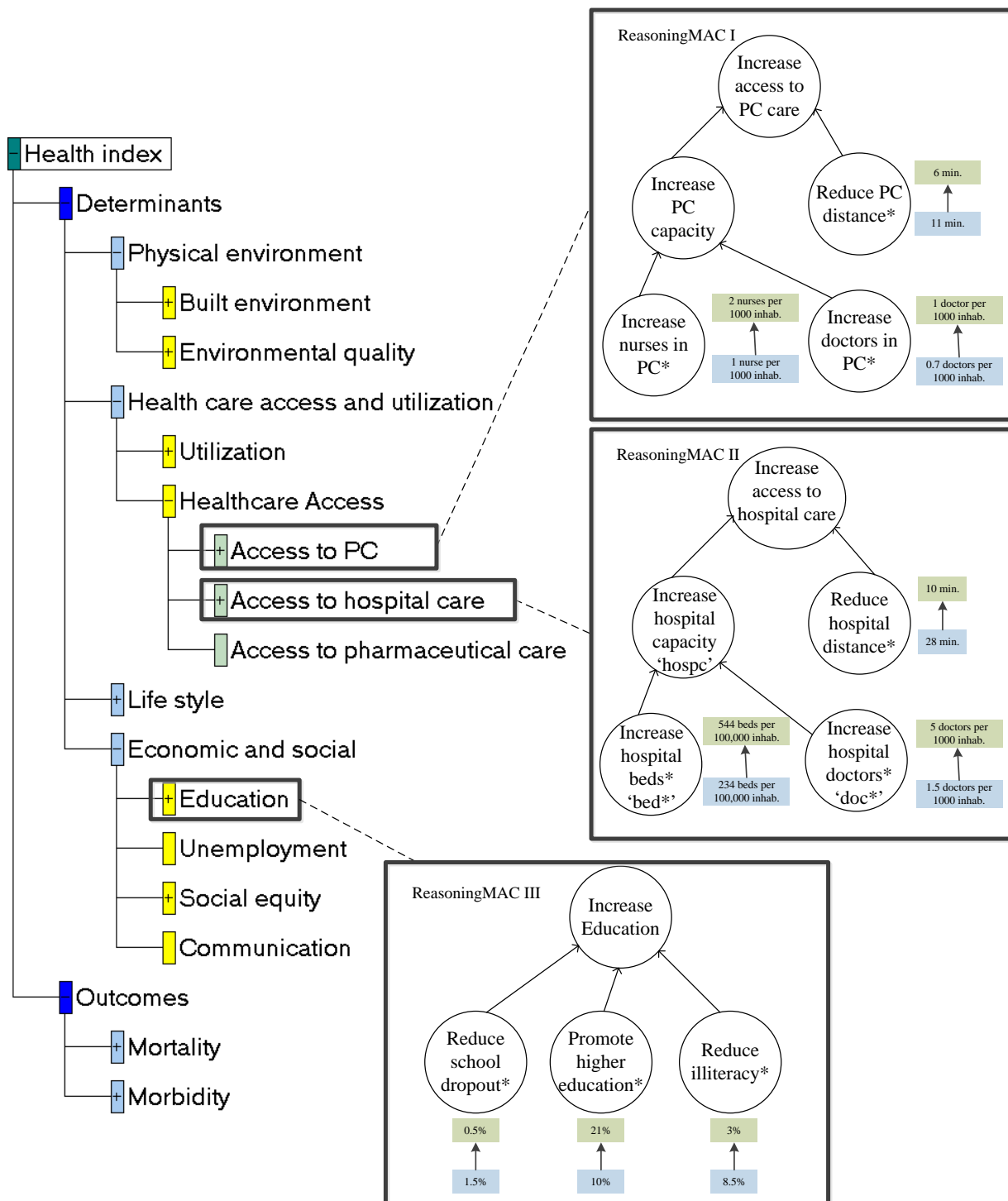


Fig. 7 – Partial view of the value tree, where the areas of concern and three of the means–ends network are depicted and where the ReasoningMAC method was applied. Source: <http://www.uc.pt/fluc/gigs/GeoHealthS>.

4.2.2. Using the ReasoningMAC method to support the development of a population health index

Once the structure of the network was agreed, we carried out qualitative tests, with the help of the M-MACBETH DSS, to verify the difference independent condition between two nodes of the network (see Figure 7), employing the *good* and *neutral* levels and the test suggested in section 3.3.4.

Figure 8 provides representations that capture the process used for testing the influence of the concepts *Hospital doctors** (*doc**) and *Hospital beds** (*bed**) on the parent node *Hospital capacity*

(*hospc*). Consider that $G_{doc^*_{hospc}}$ and $N_{doc^*_{hospc}}$ are the *good* and *neutral* levels of *doc**, respectively; and $G_{bed^*_{hospc}}$ and $N_{bed^*_{hospc}}$ the *good* and *neutral* levels of *bed**. To test that the nodes *doc** and *bed** are difference independent, on the *parent* node *hospc*: the difference in value between the performance pairs $[G_{doc^*_{hospc}}, G_{bed^*_{hospc}}]$ and $[G_{doc^*_{hospc}}, N_{bed^*_{hospc}}]$ is equal to the difference in value between the pairs $[N_{doc^*_{hospc}}, G_{bed^*_{hospc}}]$ and $[N_{doc^*_{hospc}}, N_{bed^*_{hospc}}]$; and the difference in value between the pairs $[G_{doc^*_{hospc}}, G_{bed^*_{hospc}}]$ and $[N_{doc^*_{hospc}}, G_{bed^*_{hospc}}]$ is equal to the difference in value between the pairs $[G_{doc^*_{hospc}}, N_{bed^*_{hospc}}]$ and $[N_{doc^*_{hospc}}, N_{bed^*_{hospc}}]$. Mathematically, this implies that the following conditions are included within M-MACBETH:

$$\frac{v_{hospc}(G_{doc^*_{hospc}}, G_{bed^*_{hospc}}) - v_{hospc}(N_{doc^*_{hospc}}, G_{bed^*_{hospc}})}{v_{hospc}(G_{doc^*_{hospc}}, N_{bed^*_{hospc}}) - v_{hospc}(N_{doc^*_{hospc}}, N_{bed^*_{hospc}})} = 1 \quad (7)$$

$$\frac{v_{hospc}(G_{doc^*_{hospc}}, G_{bed^*_{hospc}}) - v_{hospc}(G_{doc^*_{hospc}}, N_{bed^*_{hospc}})}{v_{hospc}(N_{doc^*_{hospc}}, G_{bed^*_{hospc}}) - v_{hospc}(N_{doc^*_{hospc}}, N_{bed^*_{hospc}})} = 1 \quad (8)$$

Three steps were carried out for this test. Firstly, qualitative judgements for the difference in value between four hypothetical options were asked to our DM group (see Figure 8.a). The ‘group judgement’ answers – reflecting the opinion of the group of DMs – were recorded in the M-MACBETH matrix from Figure 8.b. Secondly, a set of constraints that represent conditions of additivity (Eqs. 7 and 8) were introduced into the M-MACBETH software. Thirdly, M-MACBETH was used to analyse whether the matrix of judgements was compatible with the additivity condition (see Figure 8.b). In case of no compatibility – as in our reported case – the M-MACBETH matrix will be non-symmetric, implying that both concepts cannot be considered as evaluation criterion; in cases of compatibility, the MACBETH matrix will be symmetric and both concepts will qualify as candidates to be evaluation criteria. This qualification depends on the result of the following additivity test – continuing the analysis of the network II from Figure 7, now the additivity test was applied between the concepts *Hospital capacity* and *Hospital proximity* and the additivity condition was verified. Given these results, *Hospital proximity* and *Hospital capacity* will integrate the value tree as evaluation criteria within the area of concern *Healthcare access*. *Hospital capacity* will be a criterion whose value is a function of the value of the concepts *Hospital doctors* and *Hospital beds*, and measures the extent to which the hospital capacity is responding to inpatient needs.

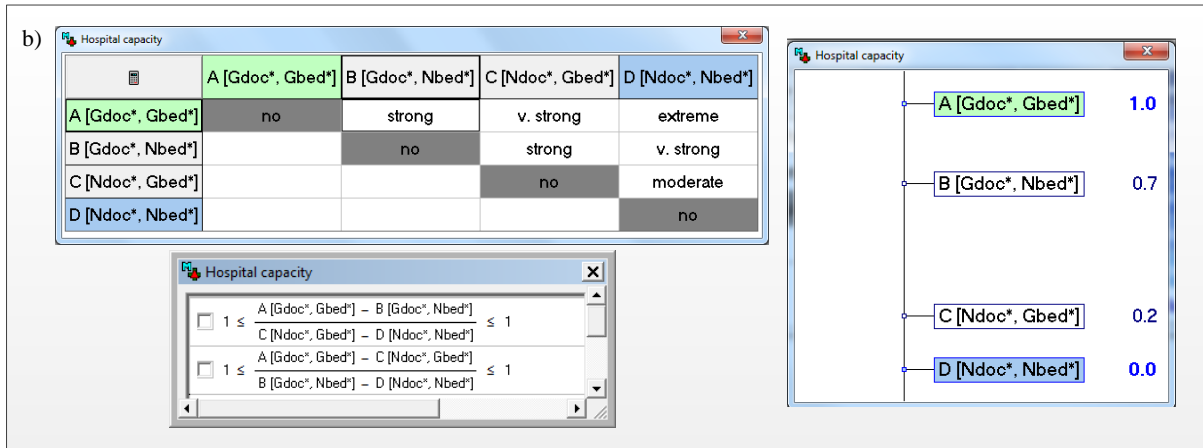
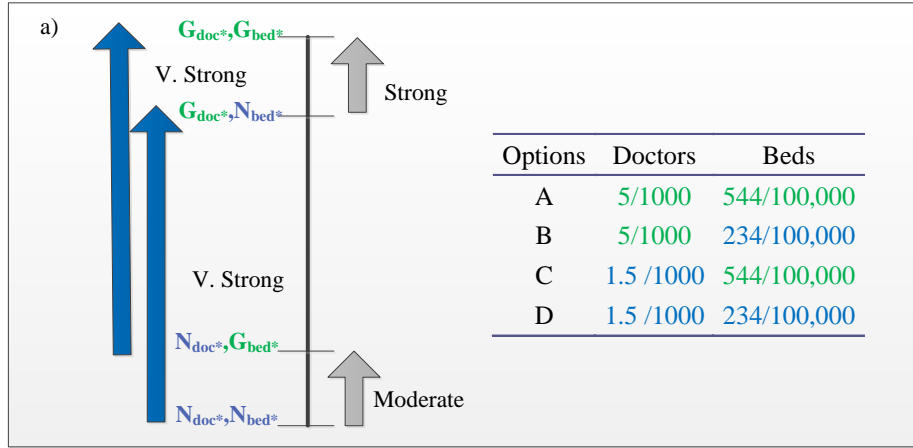


Fig. 8 - Testing the additivity condition between the *hospital doctors* (doc^*) and *hospital beds* (bed^*) concepts with the ReasoningMAC methods: a) options and protocol of questions used to test the additivity condition; b) M-MACBETH DSS employed to support the additivity test (note that the MACBETH matrix is non-symmetric against its main diagonal, which implies that both concepts are preference dependent).

To build the value function for the *parent* node *Hospital capacity*, the following steps were then performed: the judgments shown in the matrix of Figure 8.b were adjusted and validated by the DM group, leading to the options' scores depicted on the thermometer of the same figure. Then, by applying the multilinear value model it was possible to assess the weight of each concept plus the synergic effect between them. In this case, 0.70 and 0.20 were the weights of the concepts *Hospital doctors* and *Hospital beds*, respectively, and a synergic effect of 0.10 was obtained. Hence, the value (function) of the *parent* node *Hospital capacity* is given by Eqs. 4-6, where $\beta_{hospc}(x_{doc^*_{hospc}}, x_{bed^*_{hospc}})$ is modelled by equation (9) and with

$$Y_{hospc} = \left\{ v_{hospc}(x_{doc^*_{hospc}}^0, x_{bed^*_{hospc}}), v_{hospc}(x_{doc^*_{hospc}}, x_{bed^*_{hospc}}^0) \right\}.$$

$$\beta_{hospc}(x_{doc^*_{hospc}}, x_{bed^*_{hospc}}) = 0.70 v_{doc^*_{hospc}}(x_{doc^*_{hospc}}) + 0.20 v_{bed^*_{hospc}}(x_{bed^*_{hospc}}) + (-1)^l(0.10) \left| v_{doc^*_{hospc}}(x_{doc^*_{hospc}}) \right| \cdot \left| v_{bed^*_{hospc}}(x_{bed^*_{hospc}}) \right| \quad (9)$$

Eqs. 4-6 depicts the final value of an option on the parent node *Hospital capacity*, which is given by the multilinear model, while the value of an option on the parent node *Access to hospital care* is given by the simple additive model (given difference independence between *Hospital capacity* and *Hospital proximity*).

In our case study, once the ReasoningMAC method was used in the networks presented in Figure 7, it was straightforward to verify that only the concepts *Doctors* and *Nurses in primary care* (depicted in ReasoningMAC I of Figure 7) and the concepts *Hospital doctors* and *Hospital beds* (depicted in ReasoningMAC II of Figure 7) did not verify the additivity condition and consequently two criteria (*Hospital capacity* and *PC capacity*) were then modelled through the multilinear value model (Eqs. 4-6).

4.2.3. Comparing the ReasoningMAC method with a constructed descriptor

We now compare the approach adopted in the case study, described in the section above, with a common approach in dealing with value interactions: building up a constructed descriptor (or attribute), as suggested by (Keeney, 1992), and often employed in practice – see for instance, (Bana e Costa & Beinat, 2005) and (Del Rio Vilas et al., 2013). In such approach a single descriptor is developed to assess a single value function for the combination of levels between the decomposed criteria.

Consider the above example regarding the *Hospital capacity* criterion that has associated two concepts (*doc** and *bed**) each one measured by a performance indicator (*number of doctors per 1000 inhab.* and *number of beds per 100,000 inhab.*, respectively) that were discretised, following Keeney (1992)'s approach, into four levels: $B_{doc_{hospc}^*}$, $G_{doc_{hospc}^*}$, $N_{doc_{hospc}^*}$ and $W_{doc_{hospc}^*}$ are the *best*, *good*, *neutral* and *worst* levels of *doc**; and $B_{bed_{hospc}^*}$, $G_{bed_{hospc}^*}$, $N_{bed_{hospc}^*}$ and $W_{bed_{hospc}^*}$ are the *best*, *good*, *neutral* and *worst* levels of *bed**. Consider also the following hypothetical options:
 $E = [W_{doc_{hospc}^*}, B_{bed_{hospc}^*}]$, $F = [W_{doc_{hospc}^*}, G_{bed_{hospc}^*}]$, $H = [W_{doc_{hospc}^*}, N_{bed_{hospc}^*}]$,
 $R = [W_{doc_{hospc}^*}, W_{bed_{hospc}^*}]$, $S = [N_{doc_{hospc}^*}, W_{bed_{hospc}^*}]$, $T = [G_{doc_{hospc}^*}, W_{bed_{hospc}^*}]$ and
 $Z = [B_{doc_{hospc}^*}, W_{bed_{hospc}^*}]$.

The possible performance space of $v_{hospc}(\cdot)$ is illustrated in Figure 9: clearly, the *best* and *worst* hospital capacity is described by a *best* and *worst* performances on both indicators; implying that other possible combinations are between these two bounds. To build a constructed scale in this approach the group has to provide (at least) tree ordinal judgements, resulting from comparing options E and Z, F and T and finally H and S; plus, a minimum of two cardinal judgements resulting from pairwise comparing the difference in attractiveness between E and F and F and H; and between F and H and H and R. This process would enable the assessment of the indifference curves depicted in Figure 9; implying that the constructed scale would only give an aggregated value for specific $[x_{doc_{hospc}^*}, x_{bed_{hospc}^*}]$ points, not covering all the performance space (see Figure 9).

To increase the granularity of the constructed scale more levels of the indicators would need to be considered, thus increasing the number of judgments (for instance, with 5 performance levels it would be necessary to ask for 7 judgements); and, yet, we would never have a complete set of indifference curves that covered all performance space. This implies that, additional value judgments must be elicited when assessing the value of options that are between the indifference curves – see details in (Keeney, 1992). In our case study, as shown in Figure 9 this would imply (at least) more 26 judgements (depicted by the 26 grey points, where each point represents a performance profile of a municipality under evaluation).

Thus following Keeney (1992)’s approach, one needs to assess a minimum of $(n - 1)(m - 1) + (m - 2)$ judgements (where n is the number of evaluation aspects and respective indicators and m the number of performance levels of the indicator). In addition, one needs to add the number of judgements required to assess the values scores of all options that have a performance profile that is between the assessed indifference curves.

With the ReasoningMAC method instead, which requires a minimum of 3 judgements (up to the total recommended number of 6 judgements), we can develop a value scale for the *Hospital capacity* criterion that has all the performance space covered. Therefore it makes it possible to assess an aggregated value for all $[x_{doc_{hospc}^*}, x_{bed_{hospc}^*}]$ points. This brief comparison hopefully shows the potential added value of using the ReasoningMAC method for assessing a value scale for a criterion that encompasses several value interrelated health aspects. In addition, the case-study demonstrated the feasibility of modelling and eliciting the parameters required by the method.

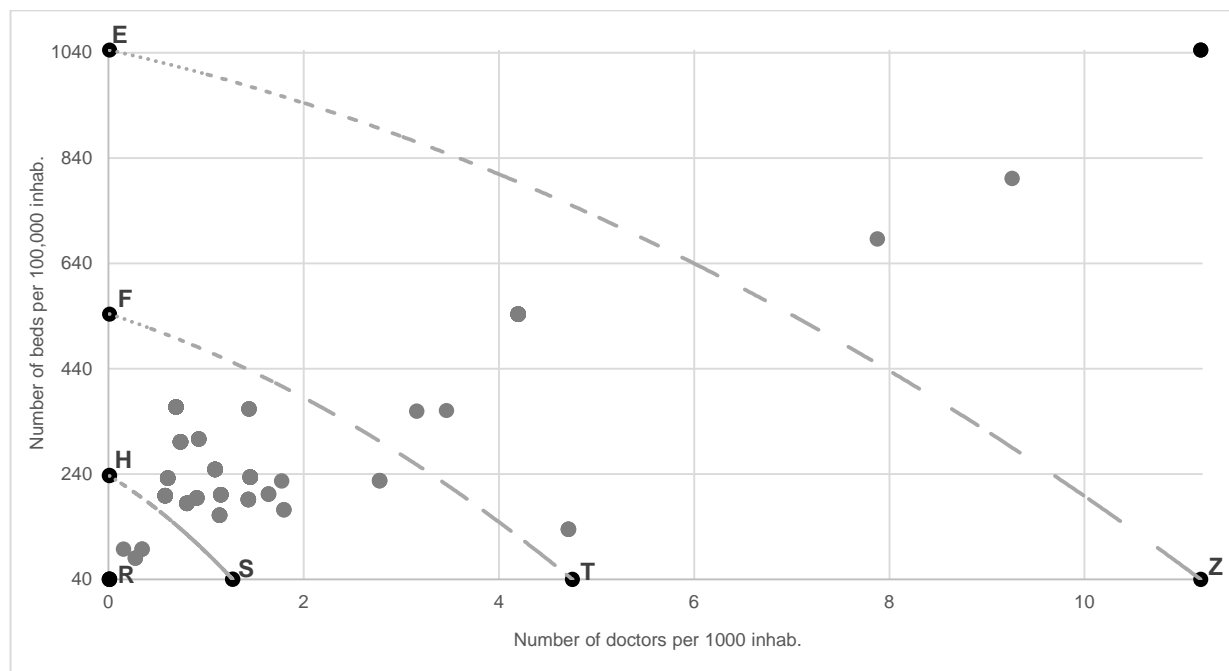


Fig. 9 – Assessing a value scale for the *Hospital capacity* criterion following the approach proposed in (Keeney, 1992), where: the dash lines represent indifference curves; and the grey points represent municipalities in which extra value judgement would need to be elicited. Notice that, despite having 278 municipalities in the evaluation set, we only show 26 grey dots, as several municipalities have the same performance profile.

5. Discussion and conclusions

Idiographic cognitive maps (ICMs) have been extensively and successfully employed in Operational Research to support problem structuring for individual decision-makers and groups (Eden, 1994; Eden & Ackermann, 2001, 2013b; Franco & Lord, 2011). Since the 1990s there have been proposals to integrate ICMs with multicriteria analysis, following the pioneer paper by (Belton et al., 1997). Nevertheless, these proposals have not fully explored the richness that a causal map has. More recently some approaches have suggested employing the structure of the map itself to assess options, among them the Reasoning Maps method (Montibeller et al., 2008). The method was designed as a decision aiding technique to support DMs during both problem structuring and the evaluation of options, integrating both into a single framework and thus avoiding the loss of the rich structure of value interactions that exist along the means-end chains. However, and despite its potential, this method faced indetermination and indistinction problems, neither modelling explicitly value interactions among evaluation concepts nor enabling a quantitative rank of options.

To overcome these limitations, we proposed extending the Reasoning Maps method and combining it with the MACBETH approach, which we denominated as the Reasoning-MACBETH map (ReasoningMAC). MACBETH supports the measurement of the strength of influence links and provides quantitative value assessments, while maintaining the key feature of the Reasoning Maps method of only asking DMs for qualitative information. ReasoningMACs are aligned with the suggestion of (Phillips, 1989, p. 89) in that ‘words are essential, more essential than numbers, but a blending of the two can enable individuals and groups to achieve new depths of understanding which would not have been possible using either words or numbers alone’.

In our view, the ReasoningMAC method is a potential alternative or a complement to other methods employed to build multicriteria value models. It is an alternative when the conditions of a simple additive model cannot be assumed, and a network of means-ends concepts is useful to evaluate options. In these cases, ReasoningMACs intend to support the evaluation of options after a problem structuring phase in which an ICM has been developed. Following this approach typically requires simplifying the rich ICM – which has a large number of links and nodes and complex relationships –, towards a model suitable for a (relatively) rich evaluation of options using the structure of the map (for instance, when compared with standard multicriteria value models). On the other hand, ReasoningMACs may be used in combination with other evaluation models when value interactions are only present between a confined set of concepts (for instance, within a large value tree), with ReasoningMACs being able to model explicitly those interactions and generating inputs for a simple additive evaluation model. Note that an implicit or explicit testing for difference independence is a necessary step on the process of building multicriteria evaluation models (Kirkwood, 1997). Even the authors who use simple additive models implicitly acknowledge the need for testing additivity conditions – in fact, in several contexts the condition of additivity is taken as a working hypothesis in the process of defining criteria and

constructing descriptors and value functions – see, for instance (Bana e Costa et al., 2012a) and (Quintino et al., 2015).

As described in the paper in the case of development of a population health index, the use of the ReasoningMAC technique supported the construction of a multicriteria model to evaluate population health by enabling the value assessment of a *parent* node that comprises several difference dependent aspects. Three main reflections emerge from the use of the ReasoningMAC approach in this case-study. First, at the beginning of the process of construction of the index, ReasoningMACs were helpful for the DMs to understand the interrelationships between health indicators. The simple protocol of questions enabled to clarify how (population health) aspects were interrelated and if difference independence was a reasonable assumption; and it made clear for the DM group the distinction between difference dependent vs. environmental (that is physical or statistical) relationships (Bana e Costa & Beinart, 2005). Secondly, for the instances where difference independence did not hold, a constructed descriptor was not an option since it would require a higher number of judgements (as was shown in section 4.2.3). Employing ReasoningMACs was thus a good compromise between simplicity and the need for quantitatively modelling the complex relations between several health aspects. Thirdly, although we did not describe in detail this aspect in this article, the social process involved in the application of the ReasoningMAC technique contributed to a higher acceptance and understanding of the ReasoningMAC outputs by the DM group, as each member of the group was able to air its concerns and views regarding population health, while analysing indicators and their value interactions; our effort to apply the ReasoningMAC method on that context was thus rewarding. In addition, the use of MACBETH questioning protocol enabled the elicitation of qualitative preference information, which might be more suitable for DMs than numerical modes of preference elicitation (Fasolo and Bana e Costa, 2014).

Regardless of these benefits, the development and application of the ReasoningMAC approach has raised several theoretical and practical questions, indicating areas for future research. Some limitations and areas identified as relevant for further research are:

- *Static analysis*: In the same way as the original Reasoning Maps method, a ReasoningMAC is designed to be used in acyclical network, performing a static analysis (as it is common in MCDA evaluations). If loops are important and are not a result of a coding mistake (Eden & Ackermann, 1992), other methods should be used to analyse the dynamic behaviour of the system. This opens a line for future research, where our approach can be further developed with the aim of modelling loops and for performing a dynamic analysis.
- *Size of the model*: a ReasoningMAC has a collapsed and simplified structure, being relatively small and less rich when compared with ICMs (Eden, 2004). Specifically, the ReasoningMAC method is designed to be used within a ‘requisite’ network in the sense that the concepts and connections depicted are enough to perform a multicriteria evaluation. Therefore, if the network cannot be reduced to a simpler and collapsed representation, this technique may not be a valid choice to perform a multicriteria evaluation of options. In line with this limitation, two perspectives should

be further investigated. Firstly, future research is needed for proposing and developing guidelines that help facilitators in building collapsed and simplified maps that can be used in a ReasoningMAC framework. Until now the process of converting a large map into a collapsed map is based on the experience and know-how of facilitators who have been employing ICMs followed by multicriteria value models. Secondly, future research could be conducted to develop further the ReasoningMAC approach for larger maps using less demanding preference elicitation protocols, e.g. eliciting only ordinal pairwise comparisons (as for example in PAPRIKA (Hansen & Ombler, 2008)).

- *Building up the model:* As the use of the ReasoningMAC approach is time consuming, it only makes sense to use it when simpler models (for instance, a simple additive multicriteria model) cannot be applied. This is compounded by the lack of specific software to support the process. Future research is needed for designing a software system which can enhance the applicability of this technique by increasing the transparency of the process and facilitating a more in-depth analysis of the ReasoningMAC results.
- *Gain-loss bias:* The ReasoningMAC method assumes that the behaviour of DM when faced with ‘positive’ and ‘negative’ scores is the same. This assumption is considered when the multilinear model is generalised to be used with bipolar scales and when negative arrows are converted into positive ones. This is not an exclusive assumption of our model, as it is often considered as a working hypothesis in the process of building multicriteria models – in fact, methods that do not entail it, such as those working with bi-capacities (Grabisch & Labreuche, 2005a, 2005b), are prohibitive in many real-world applications given their elicitation burden. Nevertheless, we are aware that experts and DMs are subject to the ‘gain-loss’ bias i.e. often behave differently in face of gains or losses (Kahneman & Tversky, 1979). Therefore, counter measures, as the use of group processes and cross-checking influence links with trade-off judgements, need to be considered during the application of our method – suggestions on how to de-bias the ‘gain-loss’ bias can be found in (Montibeller & von Winterfeldt, 2015).
- *Converting negative arrows into positive ones:* To avoid the problem of indetermination, the ReasoningMAC method does not have negative arrows, implying that dilemmas and other important relationships are not directly represented in the network. However, the method acknowledges the importance of dilemmas in decision making, by modelling them quantitatively as value trade-offs. This elimination of negative arrows also implies that our network moves away from the *thinking* of the DM originally represented in the ICM (which can be regarded as a cost of the method). Further research is needed to explore how other techniques/procedures – such as the Evaporating Cloud procedure (Goldratt, 1990, 1994) that helps DMs to understand and recognise underlying assumptions involved in dilemmas – can be used to facilitate the process of modelling dilemmas.

- *Weak difference independence*: The condition of weak difference independence may not be fulfilled in every evaluation, which will deem the use of a multilinear value model unfeasible in certain contexts. In those cases, a direction of further research is exploring the application of ReasoningMAC with the Choquet Integral operator, in particular with the MACBETH-Choquet global matrix (Lopes et al., 2014), which would require weaker assumptions.
- *Testing the usability (comparison with other methods)*: The ReasoningMAC approach was designed to evaluate options when there are value interactions between aspects that need to be addressed, and a simple additive framework cannot be used. Moreover, the approach may support the construction of simple additive models, when a (confined set of) difference dependent aspects exist and building a constructed descriptor is not an option (as exemplified in our case study). Nevertheless, and despite its potential, the usefulness of the proposed method is yet to be assessed. For this assessment an experimental design would need to be conducted with the aim of comparing in full ReasoningMAC with other methods/approaches that can also deal with multicriteria value interactions and/or use a cognitive map to perform a multicriteria evaluation. Dimensions that are relevant are, for instance, the number and nature of preference information needed, the complexity of the network and method's underlying assumptions. In addition, applying the ReasoningMAC method into other field would provide a better understanding of its *pros* and *cons*. We suggest three areas where problem structuring is required and value interactions are expected, indicating potential fields where ReasoningMAC method may be applied. Firstly in *Risk management*, as ReasoningMACs may help risk managers to take into consideration value interaction between risks impacts, which are often neglected, being an alternative when weak difference independence hold to the recently proposed MACBETH-Choquet approach (Lopes et al., 2014). Secondly in *Performance monitoring and evaluation*, as often the performance of a system is dependent on the interaction of several aspects and factors (Bertalanffy, 1969; Senge, 1990). Hence, as the ReasoningMAC method models value interactions between network nodes, it may support DMs in this field by analysing how an improvement of performance on a *descriptor* node can contribute to an increase of the added value on a *parent* node. Thirdly, in *Designing scenarios*, as ReasoningMACs can be used to support the construction of scenarios, by permitting to evaluate the value score that different options (with different performances on the *descriptor* nodes) have in the *head* nodes of the network. This information can be valuable to design contingent solutions.

Concluding, while recognising that the method does have some limitations and thus may benefit from further research and development, we hope it provides a worthwhile contribution to the relevant problem of value interactions and to the thriving use of idiographic causal maps to evaluate the multicriteria value of decision options.

Acknowledgements

The first author acknowledges financing from the Fundação para a Ciência e Tecnologia (Portugal) within Doctoral Fellowship SFRH/BD/75916/2011 and project GeoHealthS PTDC/CS-GEO/122566/2010. The third and fourth authors acknowledge funding from Fundação para a Ciência e a Tecnologia, within the project IRIS – Project risk management: Improving Risk matrices using multiple criteria decision analysis (PTDC/EGE-GES/119230/2010).

References

- Ackermann, F. (2012). Problem structuring methods ‘in the Dock’: Arguing the case for Soft OR. *Eur J Oper Res*, 219(3), 652-658.
- Ackermann, F., Howick, S., Quigley, J., Walls, L., & Houghton, T. (2014). Systemic risk elicitation: Using causal maps to engage stakeholders and build a comprehensive view of risks. *Eur J Oper Res*, 238(1), 290-299.
- Angelis, A., & Kanavos, P. (2016). Value-based assessment of new medical technologies: Towards a robust methodological framework for the application of multiple criteria decision analysis in the context of health technology assessment. *Pharmacoeconomics*, 34(5), 435-446.
- Azadeh, A., Salehi, V., Arvan, M., & Dolatkah, M. (2014). Assessment of resilience engineering factors in high-risk environments by fuzzy cognitive maps: A petrochemical plant. *Saf Sci*, 68, 99-107.
- Azadeh, A., Zarrin, M., Abdollahi, M., Noury, S., & Farahmand, S. (2015). Leanness assessment and optimization by fuzzy cognitive map and multivariate analysis. *Expert Syst Appl*, 42(15-16), 6050-6064.
- Bana e Costa, C. A., & Beinat, E. (2005). *Model-structuring in public decision-aiding*. Working paper LSE OR 05.79. London School of Economics and Political Science.
- Bana e Costa, C. A., Carnero, M. C., & Oliveira, M. D. (2012a). A multi-criteria model for auditing a Predictive Maintenance Program. *Eur J Oper Res*, 217(2), 381-393.
- Bana e Costa, C. A., & Chagas, M. P. (2004). A career choice problem: an example of how to use MACBETH to build a quantitative value model based on qualitative value judgments. *Eur J Oper Res*, 153(2), 323-331.
- Bana e Costa, C. A., Corrêa, E. C., De Corte, J.-M., & Vansnick, J.-C. (2002). Facilitating bid evaluation in public call for tenders: a socio-technical approach. *Omega- Int J Manage Sci*, 30(3), 227-242.
- Bana e Costa, C. A., De Corte, J.-M., & Vansnick, J.-C. (2012b). MACBETH. *Int J Inf Tech Decis*, 11(02), 359-387.
- Bana e Costa, C. A., De Corte, J.-M., & Vansnick, J.-C. (2015). M-MACBETH version 3.0.0 (beta) user's guide. http://www.m-macbeth.com/en/download_EN.aspx.
- Bana e Costa, C. A., De Corte, J.-M., & Vansnick, J.-C. (2016). On the mathematical foundations of MACBETH. In S. Greco, M. Ehrgott & R. J. Figueira (Eds.), *Multiple Criteria Decision Analysis: State of the art surveys* (2nd ed., pp. 421-463). New York, NY: Springer New York.
- Bana e Costa, C. A., Ensslin, L., Cornêa, É. C., & Vansnick, J.-C. (1999). Decision support systems in action: integrated application in a multicriteria decision aid process. *Eur J Oper Res*, 113(2), 315-335.
- Bana e Costa, C. A., Lourenço, J. C., Oliveira, M. D., & Bana e Costa, J. C. (2014). A socio-technical approach for group decision support in public strategic planning: The Pernambuco PPA case. *Group Decis Negot*, 23(1), 5-29.
- Bana e Costa, C. A., & Oliveira, M. D. (2012). A multicriteria decision analysis model for faculty evaluation. *Omega-Int J Manage Sci*, 40(4), 424-436.
- Bana e Costa, C. A., & Pirlot, M. (1997). Thoughts on the future of the multicriteria field: Basic convictions and outline for a general methodology. In J. Clímaco (Ed.), *Multicriteria analysis* (pp. 562-568). Berlin: Springer.
- Bana e Costa, C. A., & Vansnick, J.-C. (1994). MACBETH - An interactive path towards the construction of cardinal value functions. *Int T Oper Res*, 1(4), 489-500.
- Bana e Costa, C. A., & Vansnick, J. C. (2008). A critical analysis of the eigenvalue method used to derive priorities in AHP. *Eur J Oper Res*, 187(3), 1422-1428.
- Baykasoğlu, A., & Gölcük, İ. (2015). Development of a novel multiple-attribute decision making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS. *Inform Sciences*, 301, 75-98.
- Belton, V., Ackermann, F., & Shepherd, I. (1997). Integrated support from problem structuring through to alternative evaluation using COPE and V.I.S.A. *J Multi-Crit Decis Anal*, 6(3), 115-130.
- Belton, V., & Stewart, T. (2010). Problem structuring and multiple criteria decision analysis. In M. Ehrgott, J. R. Figueira & S. Greco (Eds.), *Trends in Multiple Criteria Decision Analysis* (pp. 209-239). New York, NY: Springer.
- Belton, V., & Stewart, T. J. (2002). *Multiple criteria decision analysis: an integrated approach*. Dordrecht: Springer.
- Bertalanffy, L. V. (1969). *General System Theory: Foundations, development, applications*. New York, NY: George Braziller.
- Bier, V. M. (2001). On the state of the art: risk communication to the public. *Reliab Eng Syst Safe*, 71(2), 139-150.
- Brosig, J., Traulsen, I., & Krieter, J. (2016). Multicriteria evaluation of classical swine fever control strategies using the Choquet integral. *Transbound Emerg Dis*, 63(1), 68-78.
- Budescu, D. V., Weinberg, S., & Wallsten, T. S. (1988). Decisions based on numerically and verbally expressed uncertainties. *J Exp Psychol Hum Percept Perform*, 14(2), 281-294.
- Clivillé, V., Berrah, L., & Mauris, G. (2007). Quantitative expression and aggregation of performance measurements based on the MACBETH multi-criteria method. *Int J Prod Econ*, 105(1), 171-189.

- Currim, I. S., & Sarin, R. K. (1984). A comparative evaluation of multiattribute consumer preference models. *Manag Sci*, 30(5), 543-561.
- Del Rio Vilas, V. J., Voller, F., Montibeller, G., Franco, L. A., Sribhashyam, S., Watson, E., et al. (2013). An integrated process and management tools for ranking multiple emerging threats to animal health. *Prev Vet Med*, 108(2-3), 94-102.
- Dyer, J., & Lorber, H. (1982). The multiattribute evaluation of program-planning contractors. *Omega- Int J Manage Sci*, 10(6), 673-678.
- Dyer, J. S. (1990). Remarks on the analytic hierarchy process. *Manag Sci*, 36(3), 249-258.
- Dyer, J. S., & Sarin, R. K. (1979). Measurable multiattribute value functions. *Oper Res*, 27(4), 810-822.
- Eden, C. (1992). On the nature of cognitive maps. *J Manag Stud*, 29(3), 261-265.
- Eden, C. (1994). Cognitive mapping and problem structuring for system dynamics model building. *Syst Dyn Rev*, 10(2-3), 257-276.
- Eden, C. (2004). Analyzing cognitive maps to help structure issues or problems. *Eur J Oper Res*, 159(3), 673-686.
- Eden, C., & Ackermann, F. (1992). The analysis of cause maps. *J Manag Stud*, 29(3), 309-324.
- Eden, C., & Ackermann, F. (1998a). Analysing and comparing idiographic causal maps. In C. Eden & J.-C. Spender (Eds.), *Managerial and organizational cognition: Theory, methods and research* (pp. 192-209). London: Sage Publications.
- Eden, C., & Ackermann, F. (1998b). *Making strategy: The journey of strategic management*. London: Sage Publications.
- Eden, C., & Ackermann, F. (2001). Soda - The principles. In J. Rosenhead & J. Mingers (Eds.), *Rational analysis for a problematic world revisited - Problem structuring methods for complexity, uncertainty and conflict* (pp. 21-41). Chichester: John Wiley & Sons.
- Eden, C., & Ackermann, F. (2004). Cognitive mapping expert views for policy analysis in the public sector. *Eur J Oper Res*, 152(3), 615-630.
- Eden, C., & Ackermann, F. (2013a). 'Joined-Up' policy-making: Group decision and negotiation practice. *Group Decis Negot*, 23(6), 1385-1401.
- Eden, C., & Ackermann, F. (2013b). Problem structuring: on the nature of, and reaching agreement about, goals. *Eur J Decis Process*, 1(1), 7-28.
- Elomda, B. M., Hefny, H. A., & Hassan, H. A. (2013). An extension of fuzzy decision maps for multi-criteria decision-making. *Egypt Inform J*, 14(2), 147-155.
- Ensslin, L., Dutra, A., & Ensslin, S. R. (2000). MCDA: a constructivist approach to the management of human resources at a governmental agency. *Int T Oper Res*, 7(1), 79-100.
- Fasolo, B., & Bana e Costa, C. A. (2014). Tailoring value elicitation to decision makers' numeracy and fluency: Expressing value judgments in numbers or words. *Omega-Int J Manage Sci*, 44, 83-90.
- Franco, L. A. (2013). Rethinking Soft OR interventions: Models as boundary objects. *Eur J Oper Res*, 231(3), 720-733.
- Franco, L. A., & Lord, E. (2011). Understanding multi-methodology: evaluating the perceived impact of mixing methods for group budgetary decisions. *Omega-Int J Manage Sci*, 39(3), 362-372.
- Franco, L. A., & Montibeller, G. (2010). Facilitated modelling in operational research. *Eur J Oper Res*, 205(3), 489-500.
- Franco, L. A., & Montibeller, G. (2011). Problem structuring for multicriteria decision analysis interventions. In J. J. Cochran, L. A. Cox, P. Keskinocak, J. P. Kharoufeh & J. C. Smith (Eds.), *Wiley Encyclopedia of Operations Research and Management Science*. New York, NY: John Wiley & Sons, Inc.
- French, S. (1984). Fuzzy decision analysis: Some criticisms. In H. J. Zimmermann, L. A. Zadeh, & B. R. Gaines (Eds.), *TIMS Study on Management Science* (pp. 29-44). Amsterdam: Elsevier.
- French, S. (1986). *Decision theory: an introduction to the mathematics of rationality*. Chichester: Ellis Horwood Limited.
- Froelich, W., Papageorgiou, E. I., Samarinas, M., & Skriapas, K. (2012). Application of evolutionary fuzzy cognitive maps to the long-term prediction of prostate cancer. *Appl Soft Comp*, 12(12), 3810-3817.
- Goldratt, E. M. (1990). *Theory of Constraints: What is this thing called the Theory of Constraints and how should it be implemented*. Great Barrington, MA: North River Press.
- Goldratt, E. M. (1994). *It's not luck*. Great Barrington, MA: North River Press.
- Grabisch, M., & Labreuche, C. (2005a). Bi-capacities — I: definition, Möbius transform and interaction. *Fuzzy Sets Syst*, 151(2), 211-236.
- Grabisch, M., & Labreuche, C. (2005b). Fuzzy measures and integrals in MCDA. In J. Figueira, S. Greco & M. Ehrgott (Eds.), *Multiple Criteria Decision Analysis: state of the art surveys* (pp. 563-604). Berlin: Springer.
- Grabisch, M., & Roubens, M. (2000). Application of the Choquet integral in multicriteria decision making. In M. Grabisch, T. Murofush & M. Sugeno (Eds.), *Fuzzy Measures and Integrals - Theory and Applications* (pp. 348-375). Heidelberg: Physica Verlag.
- Greco, S., Matarazzo, B., & Slowinski, R. (1999). In T. Gel, T.J. Stewart, & T. Hanne (Eds), *Multicriteria decision making. Advances in MCDM models, algorithms, theory, and applications* (pp. 397-455). New York, NY: Springer.
- Hansen, P., & Ombler, F. (2008). A new method for scoring additive multi-attribute value models using pairwise rankings of alternatives. *J Multi-Crit Decis Anal*, 15(3-4), 87-107.
- Jetter, A. J., & Kok, K. (2014). Fuzzy cognitive maps for futures studies—A methodological assessment of concepts and methods. *Futures*, 61, 45-57.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263-292.
- Keeney, R., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value tradeoffs*. New York, NY: John Wiley & Sons.
- Keeney, R. L. (1992). *Value-focused thinking: A path to creative decision making*. Cambridge, MA: Harvard University Press.
- Keeney, R. L., & Gregory, R. S. (2005). Selecting attributes to measure the achievement of objectives. *Oper Res*, 53(1), 1-11.
- Kirkwood, C. (1997). *Strategic decision making: Multiobjective decision analysis with spreadsheets*. Belmont, CA: Duxbury Press.

- Konar, A. (2007). *Cognitive engineering: a distributed approach to machine intelligence*. London: Springer Science & Business Media.
- Konar, A., & Chakraborty, U. K. (2005). Reasoning and unsupervised learning in a fuzzy cognitive map. *Inf Sci*, 170(2-4), 419-441.
- Kosko, B. (1986). Fuzzy cognitive maps. *Int J Man Mach Stud*, 24(1), 65-75.
- Kwan, T. W., & Leung, H. K. (2011). A risk management methodology for project risk dependencies. *IEEE Transactions on Software Engineering*, 37(5), 635-648.
- Larichev, O. (1992). Cognitive validity in design of decision *J Manag Technol*, 1(3), 127-138.
- Lopes, D. F., Bana e Costa, C. A., Oliveira, M. D., & Morton, A. (2014, March 6-8). Using MACBETH with the Choquet integral fundamentals to model interdependencies between elementary concerns in the context of Risk Management. *Proceedings of the 3rd International Conference on Operations Research and Enterprise Systems (ICORES)* (pp. 116-126). Angers, France.
- Lopez, C., & Salmeron, J. L. (2014). Dynamic risks modelling in ERP maintenance projects with FCM. *Inf Sci*, 256, 25-45.
- Ma, Q., & Kinderen, S. (2016). Goal-based decision making - Using goal-oriented problem structuring and evaluation visualization for Multi Criteria Decision Analysis. In M. Daneva & O. Pastor (Eds.), *Requirements Engineering: Foundation for Software Quality* (pp. 19-35). Cham: Springer International Publishing.
- Marchant, T. (1999). Cognitive maps and fuzzy implications. *Eur J Oper Res*, 114(3), 626-637.
- Michnik, J. (2013). Weighted Influence Non-linear Gauge System (WINGS)-an analysis method for the systems of interrelated components. *Eur J Oper Res*, 228(3), 536-544.
- Mingers, J. (1997). Multi-paradigm multimethodology. In J. Mingers & A. Gill (Eds.), *Multimethodology: The theory and practice of combining Management Science methodologies* (pp. 1-20). New York, NY: John Wiley & Sons.
- Montibeller, 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*. Paper presented at the Operation Research Peripatetic Postgraduate Programme (ORP3), Paris.
- Montibeller, G., & Belton, V. (2006). Causal maps and the evaluation of decision options—a review. *J Oper Res Soc*, 57(7), 779-791.
- Montibeller, G., & Belton, V. (2009). Qualitative operators for reasoning maps: Evaluating multi-criteria options with networks of reasons. *Eur J Oper Res*, 195(3), 829-840.
- Montibeller, G., Belton, V., Ackermann, F., & Ensslin, L. (2008). Reasoning maps for decision aid: an integrated approach for problem-structuring and multi-criteria evaluation. *J Oper Res Soc*, 59(5), 575-589.
- Montibeller, G., Belton, V., & Lima, M. V. A. (2007). Supporting factoring transactions in Brazil using reasoning maps: a language-based DSS for evaluating accounts receivable. *Decis Support Syst*, 42(4), 2085-2092.
- Montibeller, G., & von Winterfeldt, D. (2015). Cognitive and motivational biases in decision and risk analysis. *Risk Anal*, 35(7), 1230-1251.
- Montignac, F., Noirot, I., & Chaudourne, S. (2009). Multi-criteria evaluation of on-board hydrogen storage technologies using the MACBETH approach. *Int J Hydrogen Energ*, 34(10), 4561-4568.
- Nadkarni, S., & Shenoy, P. P. (2001). A Bayesian network approach to making inferences in causal maps. *Eur J Oper Res*, 128(3), 479-498.
- Nadkarni, S., & Shenoy, P. P. (2004). A causal mapping approach to constructing Bayesian networks. *Decis Support Syst*, 38(2), 259-281.
- Papageorgiou, E. I., & Froelich, W. (2012). Multi-step prediction of pulmonary infection with the use of evolutionary fuzzy cognitive maps. *Neurocomputing*, 92(1), 28-35.
- Pape, T. (2016). Prioritising data items for business analytics: Framework and application to human resources. *Eur J Oper Res*, 252(2), 687-698.
- Phillips, L. D. (1984). A theory of requisite decision models. *Acta Psychol*, 56(1-3), 29-48.
- Phillips, L. D. (1989). Decision Analysis in the 1990s. In A. Shahani & R. Stainton (Eds.), *Tutorial papers in Operational Research* (pp. 73-90). Birmingham: The Operational Research Society.
- Phillips, L. D., & Bana e Costa, C. A. (2007). Transparent prioritisation, budgeting and resource allocation with multi-criteria decision analysis and decision conferencing. *Ann Oper Res*, 154(1), 51-68.
- Pinar, M., Cruciani, C., Giove, S., & Sostero, M. (2014). Constructing the FEEM sustainability index: A Choquet integral application. *Ecol Indic*, 39, 189-202.
- Quintino, A., Lourenço, J. C., & Catalão-Lopes, M. (2015). Risk tolerance evaluation for an oil and gas company using a multi-criteria approach. In E. Pinson, F. Valente & B. Vitoriano (Eds.), *Operations Research and Enterprise Systems* (pp. 199-214). Cham: Springer International Publishing.
- Rescher, N. (1969). *Introduction to value theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Rodrigues, T. C. (2014). The MACBETH approach to health value measurement: Building a Population Health Index in group processes. *Procedia Technology*, 16, 1361-1366.
- Roy, B. (1993). Decision science or decision-aid science? *Eur J Oper Res*, 66(2), 184-203.
- Saaty, T. L. (2004). Decision making —The analytic hierarchy and network processes (AHP/ANP). *J Syst Sci Syst Eng*, 13(1), 1-35.
- Senge, P. (1990). *The fifth discipline: The art and practice of the learning organization*. Broadway, NY: Currency Doubleday.
- Silva, J. B., Graça Saraiva, M., Ramos, I. L., & Bernardo, F. (2013). Improving visual attractiveness to enhance city-river integration - A methodological approach for ongoing evaluation. *Plan Pract Res*, 28(2), 163-185.
- Smith, J. E., & von Winterfeldt, D. (2004). Anniversary article: decision analysis in management science. *Manag Sci*, 50(5), 561-574.
- Soetanto, R., Dainty, A. R. J., Goodier, C. I., & Austin, S. A. (2011). Unravelling the complexity of collective mental models: A method for developing and analysing scenarios in multi-organisational contexts. *Futures*, 43(8), 890-907.

- Tzeng, G.-H., Chen, W.-H., Yu, R., & Shih, M.-L. (2010). Fuzzy decision maps: a generalization of the DEMATEL methods. *Soft Comput*, 14(11), 1141-1150.
- von Winterfeldt, D., & Edwards, W. (1986). *Decision analysis and behavioral research*. Cambridge, MA: Cambridge University Press.
- Wallenius, J., Dyer, J. S., Fishburn, P. C., Steuer, R. E., Zionts, S., & Deb, K. (2008). Multiple criteria decision making, multiattribute utility theory: recent accomplishments and what lies ahead. *Manag Sci*, 54(7), 1336-1349.
- Walshe, T., & Burgman, M. (2010). A framework for assessing and managing risks posed by emerging diseases. *Risk Anal*, 30(2), 236-249.
- Wellman, M. P. (1990). Graphical inference in qualitative probabilistic networks. *Networks*, 20(5), 687-701.
- Wellman, M. P. (1994). Inference in cognitive maps. *Math Comput Simul*, 36(2), 137-148.
- Wood, M. D., Bostrom, A., Bridges, T., & Linkov, I. (2012). Cognitive mapping tools: review and risk management needs. *Risk Anal*, 32(8), 1333-1348.
- Xiao, Z., Chen, W., & Li, L. (2012). An integrated FCM and fuzzy soft set for supplier selection problem based on risk evaluation. *Appl Math Model*, 36(4), 1444-1454.
- Yu, R., & Tzeng, G. H. (2006). A soft computing method for multi-criteria decision making with dependence and feedback. *Appl Math Comput*, 180(1), 63-75.
- Zaksek, M., & Arvai, J. L. (2004). Toward improved communication about wildland fire: mental models research to identify information needs for natural resource management. *Risk Anal*, 24(6), 1503-1514.