Aggregation of dependent criteria in multicriteria decision making problems by means of capacities

Gleb Beliakov[1], Sharon Boswell[2], Thang Cao[2], Richard Dazeley[1], Vicky Mak-Hau[1], Minh-Tuan Nguyen[2], Tim Wilkin[1] and John Yearwood[1]

[1]*School of Information Technology, Deakin University, Burwood 3125, Australia*
[2]*Joint & Operations Analysis Division, Defence Science and Technology Group, Australia*

Email: gleb@deakin.edu.au

Abstract: We address the multiple criteria decision problem in which the criteria are mutually dependent, in the context of land combat vehicle selection for Australian Defense. We employ the theory of capacities to model criteria dependencies. Its challenge is exponentially growing (with the number of criteria) set of parameters that represent all criteria interactions. We outline a number of strategies to (a) simplify the construction process and reduce the number of parameters, and (b) elicit preferences from the decision makers and translate them into capacity learning problem. We present various challenges and ways to address them from the mathematical programming perspective.

Keywords: Capacity, Fuzzy measure, MCDM, Choquet integral, dependent criteria

1 INTRODUCTION

This paper proposes a number of methodologies for multiple criteria decision making problems (MCDM) with interacting criteria. The subject of study is in the context of selection of land combat vehicle (LCV) platform for Australia Defence. Each of these LCVs has a number of configurations with a number of options to choose from–these are driven by the traditional desires to have appropriate levels of survivability and lethality against threats, and modern factors, including sensors, firepower, mobility, tactics, situational awareness and communications. These configurations were evaluated in Nguyen et al. (2016); Nguyen and Cao (2017) with respect to four major groups of system capability criteria: Survivability (S), Lethality (L), Mobility (M) and C4ISR[1] (C). The decision problem is thus the selection of the "best" LCV configuration based on the criteria mentioned above. Of course, the notion of "best" is very much context and purpose dependent, but at the moment we leave this issue aside.

There are various approaches to such a MCDM problem, in particular the multiattribute utility theory (MAUT) (Keeney and Raiffa, 1976), in which the utilities of the selection criteria are combined into a single overall utility value that is used to rank the alternatives. There are two important issues which frequently fall outside traditional methods: the weighting of the criteria and their mutual dependency. The former was addressed in the AHP method (Saaty, 1992) and many of its derivatives, as well as various methods of learning the criteria importance from data. The issue of criteria interaction is even more challenging due to exponentially increasing number of ways the criteria can depend on one another. The fact that the criteria interaction need to be accounted for should be clear from the mentioned MCDM problem: for example, M, C and L affect S; and C affects M.

We employ the theory of capacities and discrete nonlinear integrals to model criteria dependency, and outline the existing techniques and their limitations, the gaps in the knowledge and the approaches to overcome the limitations and fill the gaps, leading to successful modeling MCDM problems with criteria dependencies. While at the moment we treat single LCV design selection, in the future we plan to address a much larger problem of selection multiple vehicles and interacting systems in combination, by extending the techniques outlined in this paper.

The theory of capacities (Choquet, 1954; Grabisch, 2016) is also known as the cooperative games theory, and the fuzzy measure theory (Beliakov et al., 2019), the terms we will be using interchangeably. A capacity is a set function which assigns a value not just to any individual criterion, but also to their combinations. Thus a criterion may be unimportant by itself yet gain importance in combination with other criteria. To the contrary, the value, or weight, of a group of criteria may be less than the sum of the individual criteria weights. The aggregation of the individual utilities is performed by the discrete nonlinear integrals, in particular the Choquet and Sugeno integrals with respect to capacities (Beliakov et al., 2019), which account for all criteria interactions. Importantly, in this framework the criteria dependency is treated explicitly rather than a black box, which provides opportunities to explain the rationale for the choice of alternatives to decision makers.

This paper is structured as follows. Section 2 provides preliminary definitions. Section 3 details the LCV decision problem. In Section 4 we present a decision model based on the capacities and the Choquet integral, and also outline some of the challenges. Section 5 is devoted to the gaps in the knowledge and the ways to overcome them. Section 6 concludes the finding of our study.

2 CAPACITIES

Let $N = \{1, 2, \ldots, n\}$, $n \geqslant 2$, be the set of decision criteria, $\mathcal{P}(N)$ be the power set of N, and $|S|$ be the cardinality of subset $S \subseteq N$. The definitions listed in this section can be found in references (Choquet, 1954; Beliakov et al., 2016; Grabisch and Labreuche, 2010; Grabisch, 2016; Beliakov et al., 2019)

Definition 1. A capacity on N is a set function $\mu : \mathcal{P}(N) \to [0, 1]$ such that (i) $\mu(\emptyset) = 0$, $\mu(N) = 1$; (ii) $\forall A, B \subseteq N, A \subseteq B$ implies $\mu(A) \leqslant \mu(B)$.

The quantity $\mu(A)$ is a reflection of the decision maker's perception on the importance of this criteria subset for the decision problem. The additivity of the traditional measures is replaced by a more flexible condition of monotonicity with respect to set inclusion.

Capacities are used in the definition of the discrete Choquet integral, an important aggregation function which accounts for not only input importance but their interactions (Beliakov et al., 2016; Grabisch, 1996).

---

[1]Command, Control, Communications, Computer, Intelligence, Surveillance, and Reconnaissance

Definition 2. For a given $f : N \rightarrow \Re^+$, its discrete Choquet integral $\mathcal{C}$ with respect to the capacity $\mu : \mathcal{P}(N) \rightarrow [0, 1]$ is defined as follows:

$$\mathcal{C}(f) = \sum_{i=1}^{n} (f_{(i)} - f_{(i-1)})\mu(\{(i), \ldots, (n)\}), \tag{1}$$

where $f_{(.)}$ is a non-decreasing permutation induced by $f_i, i = 1, \ldots, n$, i.e., $f_{(1)} \leqslant \ldots \leqslant f_{(n)}$, and $f_{(0)} = 0$ by convention.

The Choquet integral can be expressed as follows (Beliakov et al., 2016, Ch.4),

$$\mathcal{C}_\mu(f) = \sum_{A \subseteq N} \mu(A)(\min_{i \in A} f_i - \max_{i \in N \setminus A} f_i)^+ \tag{2}$$

The interaction of multiple decision criteria can be described by means of Shapley simultaneous interaction index or other indices, such as nonadditivity index.

Definition 3. Let μ be a capacity on N. The Shapley importance and interaction index of μ is defined as

$$I_\mu(A) = \sum_{B \subseteq N \setminus A} \frac{1}{(|N| - |A| + 1)} \binom{|N| - |A|}{|B|}^{-1} \left(\sum_{C \subseteq A} (-1)^{|A \setminus C|} \mu(C \cup B) \right), A \subseteq N.$$

Generally speaking, $I_\mu(\{i\})$, called the Shapley value, is regarded as the overall importance of the criterion i in all possible coalitions, and $I_\mu(A), |A| \geqslant 2$ is the simultaneous interaction index of the subset A.

Definition 4. A capacity μ on N is said to be $\star$-additive, if

$$\mu(A \cup B) \overset{\star}{=} \mu(A) + \mu(B), \forall A, B \subseteq N, A, B \neq \emptyset, A \cap B = \emptyset.$$

Furthermore, μ is said to be $\star$-additive within $S \subseteq N$, if

$$\mu(A \cup B) \overset{\star}{=} \mu(A) + \mu(B), \forall A, B \subseteq S, A, B \neq \emptyset, A \cap B = \emptyset.$$

where "$\overset{\star}{=}$" stands for "$= (\text{resp.} \geqslant, \leqslant, >, \text{and} <)$", "$\star$-additive" stands for "additive (resp. superadditive, subadditive, strict superadditive, and strict subadditive)".

It is a common interpretation that an additive capacity means that the decision criteria are independent; a strict superadditive (resp. strict subadditive, superadditive, and subadditive) capacity indicates that all criteria are mutually complementary (substitutive, nonsubstitutive, and noncomplementary).

Definition 5. The nonadditivity index of μ is defined as

$$n_\mu(A) = \mu(A) - \frac{1}{2^{|A|-1} - 1} \sum_{C \subset A} \mu(C), A \subseteq N.$$

Theorem 1. *(Wu and Beliakov, 2018) If a capacity μ on N is $\star$-additive, then $n_\mu(A) \overset{\star}{=} 0, \forall A \subseteq N$, $|A| \geqslant 2$. If a capacity μ on N is $\star$-additive within $S \subseteq N$, then $n_\mu(A) \overset{\star}{=} 0, \forall A \subseteq S, |A| \geqslant 2$.*

This means that nonadditivity index reflects the above common interpretation.

Let X be the set of alternatives, and $C_1, C_2, \ldots, C_n$ be the decision criteria. Each alternative $x \in X$ is evaluated with respect to each of the criteria, giving the individual utility values $u_i(x) \in [0, 1]$, for $i = 1, \ldots, n$. Other intervals, or even other types of ranges can be used, including intervals, fuzzy sets, linguistic labels, etc., which will be briefly mentioned in Section 5.

In the capacity-based MCDM models the values of the utilities u_i are combined into an overall value $U = f(u_1, \ldots, u_n)$ by using the Choquet integral, or another capacity-based integral, see Beliakov et al. (2019). For an additive capacity the Choquet integral becomes the usual weighted arithmetic mean (WAM), whereas for a symmetric capacity it becomes an ordered weighted averaging (OWA) function, which is a linear combination of order statistics (Yager, 1988).

3 LCV SELECTION PROBLEM

In the LCV selection problem mentioned in the introduction, there are four groups of criteria. In each group there are several subcriteria whose values are denoted by $u_k = w_{ij}$, $i = 1, 2, 3, 4$, $k = 1, \ldots, 17$, see Figure 1. The reason the criteria are grouped is that we use a hierarchical structure under the assumption that the criteria in each subgroup are independent, see the discussion in the next section.

The evaluations of each LCV system with respect to the criteria were obtained by extensive simulations and inferences from a Bayesian network (BN) model, as reported in Nguyen et al. (2016); Nguyen and Cao (2017). These values are scaled to $[0, 1]$ and will be used here as given.
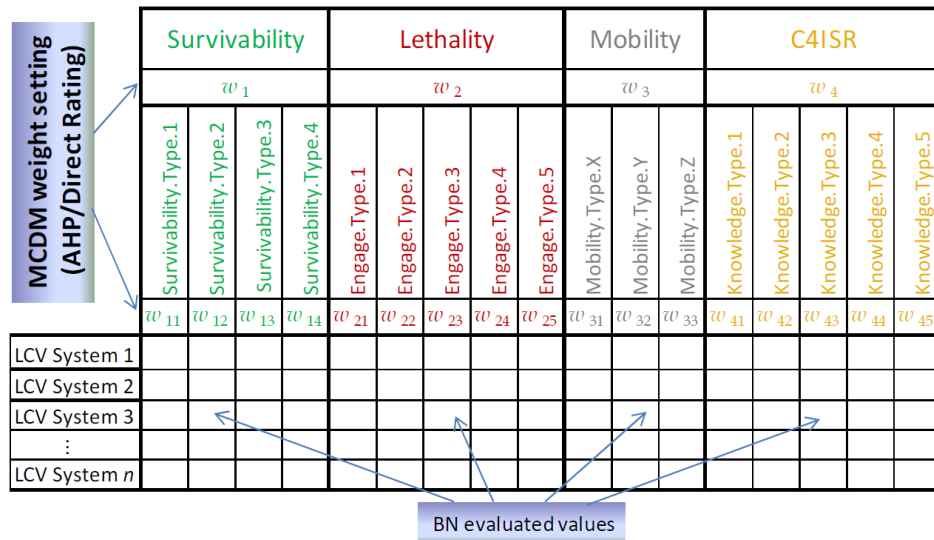

Figure 1. The evaluation criteria in the LCV problem.

There are over 15 million configurations corresponding to the rows in Figure 1.

4 MODELING DEPENDENCY THROUGH THE CHOQUET INTEGRAL

The most direct way to apply the Choquet integral is to use the formula (2) to the individual utilities u_i and selected capacity μ. The challenge is that the capacity μ is not known at this stage. There are $2^{17} - 2$ capacity values that need to be defined in order to take all criteria dependencies into account, and eliciting these values from the decision makers (DM) is undoable.

We now consider possible ways to define capacities from scarce information provided by the DMs. The first step is capacity simplification, which can take several forms. There are approaches that restrict criteria interaction to groups of at most k, called k-order capacities. The concept of k-additive fuzzy measure was developed in Grabisch (1997). Indeed, in the Möbius representation the coefficients of k-additive fuzzy measures are set to 0 whenever $|A| > k$. Another approach is to use k-maxitive fuzzy measures (Calvo and De Baets, 1998; Mesiar, 1999) which also restrict input interactions to subsets of at most k inputs. The p-symmetric capacity (Miranda et al., 2002) represents decision with p groups of anonymous referees or indifferent criteria. The case of 2-additive capacities, in which the interactions are reduced to pairs is the most widely employed.

Another way of simplifying capacity is incorporating known information about independence, which is what we plan to employ here. The groups S, L, M and C involve subcriteria w_{ij} which are deemed by the DMs as independent, therefore it makes sense to explicitly treat them as additive. One way of doing this is to set the respective interaction or nonadditivity indices $I_\mu(A)$ and $n_\mu(A)$ to zero. Since these indices are linear combinations of capacity values, that will involve a set of linear constraints, and a constraint satisfaction problem can be formulated. However this approach is rather inefficient, as we do not reduce the number of variables or monotonicity constraints.

It is feasible to incorporate partial independence information explicitly. As an example, the independence within group P implies that $\mu(\{1,2,3,4\}) = \mu(\{1\}) + \mu(\{2\}) + \mu(\{3\}) + \mu(\{4\})$, and similarly for the other groups. Then we can formulate a 4 (interacting) criteria problem and determine the capacity

$$0 = \mu(\emptyset), \mu(S), \mu(L), \mu(M), \mu(C), \mu(\{S,L\}), \ldots, \mu(\{S,L,M,C\}) = 1,$$

in which case we have only $2^4 - 2 = 14$ parameters to determine. This approach relies on so-called double aggregation functions (Calvo and Pradera, 2004), but their weights identification is an open problem.

Constructing double aggregation functions based on partial independence is not as straightforward as it looks, because the interdependency of groups of criteria implies that the criteria w_{ij} and w_{kl} from distinct groups j and l possibly interact. Therefore we potentially have many nonzero interaction indices for the capacity defined on the 17 criteria. This particular approach requires further investigation and additional simplification strategies. In particular the approach based on k-interactivity presented in Beliakov et al. (2019), could be modified by using duality of the capacities.

Another approach that is worth exploring is the p-symmetric capacities (Miranda et al., 2002), under the assumption that all the additive criteria within each group is symmetric. Here the symmetric and additive criteria are allocated to their respective equivalence classes. It is then possible to formulate the capacity learning problem in the space of equivalence classes, which we will explore in forthcoming studies.

The next issue is incorporation of other information obtainable from the DMs. This information can take the form of explicit and implicit preferences. Explicit preferences take the form of either numerical estimates, or, which is more feasible, ranking of the values of the criteria and their interactions with respect to one another. For instance the following ranking on the importance of the four criteria was obtained $S \succ L \succ M \succ C$. That can be translated into the order of the Shapley values of that capacity (rather than simply the values $\mu(S) \geqslant \mu(L) \geqslant \mu(M) \geqslant \mu(C)$).

Unfortunately it is rather hard to elicit substantially more explicit information than for individual criteria. Therefore that only gives a rather sparse set of constraints on the capacities and hence a large set of capacities compatible with the DMs explicit preferences. There are also approaches based on the DMs implicit preferences. For example, the DMs may provide their preferences on particular alternatives, or case scenarios. These preferences in the form $x \succ y$ for which the evaluation of the utilities does not result in $u_i(x) \geqslant u_i(y), \forall i = 1, \ldots, 17$ are most valuable (if the utility vectors dominate, then this is trivial information already accommodated by the monotonicity of capacities). From those non-trivial comparisons one can infer the DMs preferences about the decision criteria, and hence we call them implicit.

Learning capacities from empirical data is a known approach when the data are provided on a numerical scale (Grabisch et al., 2008; Beliakov et al., 2016; Marichal and Roubens, 2000; Beliakov, 2009; Tehrani et al., 2012). We use the pairs input–output with the prescribed values, and fit the fuzzy integral to that data, essentially in the same way the weights of an arithmetic mean can be learned from the data. However if the preferences are given on an ordinal scale, a different learning approach of the nonadditive robust ordinal regression model (NAROR) (Angilella et al., 2010; Corrente et al., 2016) needs to be employed. While it can be formulated as a linear constraints satisfaction problem, the challenge of the sparsity of these constraints remains.

Lately, a new form of preference information has been introduced and studied (Wu et al., 2015), called the MCCPI (Multiple Criteria Correlation Preference Information) which is basically a set of 2-dimensional (importance, interaction) pairwise comparison information and can be regarded as the natural generalisation of the 1-dimensional (importance) pairwise comparison information obtained by the well-known decision method, AHP (Analytic Hierarchy Process) (Saaty, 1992). The MCCPI can be obtained through an aid tool called the refined diamond diagram, which holistically composes all the situations of the overall importance and the simultaneous interaction between two criteria. With the MCCPI, some optimization models, like the least squares based nonlinear programming model and the least absolute deviation based linear programming model, are established to get the most desired 2-additive or ordinal capacity to closely caters to the DM's initial preference (Wu et al., 2015).

Thus, a combination of (a) capacity simplification and hierarchical partitioning, (b) incorporation of the explicit DM's preferences and (c) learning from the implicit DM's preference provides a viable approach to identify a set of optimal capacities that can be used to rank the alternatives through the discrete Choquet integral.

An important feature of the capacity-based MCDM model is the possibility to explain the reasoning employed in the model. Firstly, the Shapley values express the overall importance of the decision criteria in all

possible coalitions, i.e., accounting for all criteria interactions, which is a significant step forward from just the individual criteria importances (we note that Lloyd Shapley was awarded the Nobel Prize in Economics for his groundbreaking work in games theory in 2012).

Secondly, the interaction indices, as well as the nonadditivity and nonmodularity indices express the degrees of interaction and their signs, so that complementarity and redundancy of the criteria can be identified. Lastly, some input interactions can be explicitly neglected and that information is made available to the DM to assess the model meaningfulness.

5 CHALLENGES AND WAYS TO OVERCOME THEM

The great challenge of the non-additive model is its exponential complexity, i.e., the exponentially fast growth of the number of interdependencies with the number of criteria. For this reason: (a) it is not feasible to appreciate and evaluate all the dependencies by neither human decision makers nor computational models, (b) it is necessary to develop suitable mathematical simplification strategies (including partial additivity, hierarchical models, etc), (c) it is necessary to develop sophisticated high-performance computational models for learning optimal capacities.

Sparsity of the information which can be elicited from the DMs in terms of explicit or implicit preferences is another obstacle. This information does not lead to a unique capacity and hence unique ranking of the alternatives. One approach here is to determine an optimal capacity based on some additional criteria, such as the entropy (Beliakov and Wu, 2019). The higher the entropy of a capacity, the more is the (average) contribution of each individual criterion. The maximum entropy principle has been employed previously in the definition of OWA functions weights (see an overview and the references in Beliakov et al. (2019, Ch. 6)), and it makes sense to employ it for general capacities.

When the data are given on an ordinal scale, it also makes sense to look at other non-additive integrals beside the Choquet integral. There are various such integrals, and the Sugeno integral fits ordinal framework very well (Beliakov et al., 2019). The operations of sum and product are replaced with the maximum and minimum. In this case, however, fitting the capacity to the DM's preferences involves a much more challenging non-convex optimisation problem, recently discussed by Beliakov et al. (2019); Gagolewski et al. (2019). There are also case of the data given on interval scale, or as intuitionistic fuzzy sets (which are isomorphic to intervals). This type of data can be also represented as a set of linear constraints, and consequently incorporated into capacity learning problem at a reasonable price.

A different approach to optimal capacity construction is based on accommodating explicit and implicit preferences on an ordinal scale, through a set of linear constraints that restrict the set of compatible capacities. The NAROR method mentioned above is one alternative, where an additional criterion, such as entropy, can be optimised. Another approach is to sample the feasible set of capacities through a simulation, and determining the likelihood of one or another scenario. This approach requires an efficient sampling technique from a set of capacities, which is a polytope in $2^n - 2$ dimensional space. Uniform sampling from a polytope in high dimensions is another challenging problem, currently approached through random walks (such as Dikin walk). Note that it is impractical even to enumerate the vertices of the polytopes of capacities, whose number corresponds to the Dedekind numbers (only the first 8 Dedekind numbers are currently known). The efficiency of random walks is acceptable for up to 1000 variables, but this hardly sufficient for our case. Therefore new methods of uniform generation of random points in polytopes in high dimensions are needed.

6 CONCLUSIONS

We outlined an approach to tackle mutual dependencies between the decision criteria in MCDM problems, which is based on capacities, also referred to as fuzzy measures, and the discrete nonlinear integrals. The practical impact of using fuzzy measures is explicit treatment and numerical quantification of the criteria interactions, and more precise models of the decision making problem. The price for such flexibility is exponential complexity of fuzzy measures, which means limitation to 6-8 decision criteria, when no simplifications are used, resulting in 62-254 parameters. A number of strategies are needed, including capacity simplifications and their hierarchical decompositions, eliciting and representing preferences of the decision makers, mathematical optimisation and machine learning techniques to determine optimal capacities, suitable sampling techniques, and finally, methods of quantitative interpretation of the capacity values. While some capacity learning techniques can be transformed into large scale linear programming problems, others, especially the ones used for preferences or data on ordinal scales, require sophisticated nonlinear optimisation methods.

REFERENCES

Angilella, S., S. Greco, and B. Matarazzo (2010). Non-additive robust ordinal regression: A multiple criteria decision model based on the Choquet integral. *European Journal of Operational Research 201*(1), 277–288.

Beliakov, G. (2009). Construction of aggregation functions from data using linear programming. *Fuzzy Sets and Systems 160*(1), 65–75.

Beliakov, G., H. Bustince, and T. Calvo (2016). *A Practical Guide to Averaging Functions*. New York: Springer.

Beliakov, G., M. Gagolewski, and S. James (2019). Aggregation on ordinal scales with Sugeno integral for biomedical applications. *Information Sciences 501*, 377–387.

Beliakov, G., S. James, and J.-Z. Wu (2019). *Discrete Fuzzy Measures: Computational Aspects*. Berlin, Heidelberg: Springer.

Beliakov, G. and J.-Z. Wu (2019). Learning fuzzy measures from data: simplifications and optimisation strategies. *Information Sciences 494*, 100–113.

Calvo, T. and B. De Baets (1998). Aggregation operators defined by k-order additive/maxitive fuzzy measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 6*(06), 533–550.

Calvo, T. and A. Pradera (2004). Double aggregation operators. *Fuzzy Sets and Systems 142*, 15–33.

Choquet, G. (1954). Theory of capacities. *Annales de l'institut Fourier 5*, 131–295.

Corrente, S., S. Greco, and A. Ishizaka (2016). Combining analytical hierarchy process and Choquet integral within non-additive robust ordinal regression. *Omega 61*, 2–18.

Gagolewski, M., S. James, and G. Beliakov (2019). Supervised learning to aggregate data with the Sugeno integral. *IEEE Transactions on Fuzzy Systems 27*, 810–815.

Grabisch, M. (1996). The applications of fuzzy integrals in multicriteria decision making. *Europ. J. Operations Research 89*, 445–456.

Grabisch, M. (1997). k-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems 92*(2), 167–189.

Grabisch, M. (2016). *Set Functions, Games and Capacities in Decision Making*. Berlin, New York: Springer.

Grabisch, M., I. Kojadinovic, and P. Meyer (2008). A review of methods for capacity identification in Choquet integral based multi-attribute utility theory: Applications of the Kappalab R package. *European Journal of Operations Research 186*(2), 766–785.

Grabisch, M. and C. Labreuche (2010). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *Annals of Operations Research 175*(1), 247–286.

Keeney, R. and H. Raiffa (1976). *Decisions with Multiple Objectives: Performances and Value Trade-Offs*. New York: Wiley.

Marichal, J.-L. and M. Roubens (2000). Determination of weights of interacting criteria from a reference set. *European Journal of Operational Research 124*(3), 641–650.

Mesiar, R. (1999). k-order additive fuzzy measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 7*(06), 561–568.

Miranda, P., M. Grabisch, and P. Gil (2002). p-symmetric fuzzy measures. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 10*(supp01), 105–123.

Nguyen, M.-T. and T. Cao (2017). A hybrid decision making model for evaluating land combat vehicle system. In G. Syme, D. Hatton MacDonald, B. Fulton, and J. Piantadosi (Eds.), *22nd International Congress on Modelling and Simulation, MODSIM2017*, Hobart, Australia, pp. 1399–1405.

Nguyen, M.-T., T. Cao, and W. Chau (2016). Bayesian network analysis tool for land combat vehicle system evaluation. In *the 24th National Conference of the Australian Society for Operations Research (ASOR)*, Canberra, Australia.

Saaty, T. (1992). *Multicriteria Decision Making – The Analytic Hierarchy Process*. Pittsburgh: RWS Publications.

Tehrani, A., W. Cheng, K. Dembczynski, and E. Hüllermeier (2012). Learning monotone nonlinear models using the Choquet integral. *Machine Learning 89*(1-2), 183–211.

Wu, J.-Z. and G. Beliakov (2018). Nonadditivity index and capacity identification method in the context of multicriteria decision making. *Information Sciences 467*, 398–406.

Wu, J.-Z., S. Yang, Q. Zhang, and S. Ding (2015). 2-additive capacity identification methods from multi-criteria correlation preference information. *IEEE Transactions on Fuzzy Systems 23*(6), 2094–2106.

Yager, R. (1988). On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems, Man and Cybernetics 18*, 183–190.