European Journal of Operational Research 223 (2012) 1–14

Contents lists available at SciVerse ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Invited Review

Modeling uncertainty in multi-criteria decision analysis

Ian N. Durbach [a,*], Theodor J. Stewart [a,b]

[a] Department of Statistical Sciences, University of Cape Town, South Africa
[b] Manchester Business School, University of Manchester, Manchester M15 6PB, United Kingdom

ARTICLE INFO

Article history:
Received 20 January 2012
Accepted 25 April 2012
Available online 18 May 2012

Keywords:
Decision analysis
Multiple criteria analysis
Risk management
Uncertainty modeling

ABSTRACT

This paper provides a review of multiple criteria decision analysis (MCDA) for cases where attribute evaluations are uncertain. The main aim is to identify different tools which can be used to represent uncertain evaluations, and to broadly survey the available decision models that can be used to support uncertain decision making. The review includes models using probabilities or probability-like quantities; explicit risk measures such as quantiles and variances; fuzzy numbers, and scenarios. The practical assessment of uncertain outcomes and preferences associated with these outcomes is also discussed.

1. Introduction

The practice of multi-criteria decision analysis (MCDA) is concerned with the evaluation of a set of possible courses of action or alternatives. This evaluation may take a number of forms – selecting a preferred alternative, ranking alternatives from best to worst, sorting the alternatives into ordered classes such as 'good' or 'bad', among others. The set of alternatives may be either explicitly defined and discrete in number or implicitly defined via constraints in a mathematical programming formulation. It is in the former setting that greater focus is placed on explicit preference modeling, and it is here that the focus of the present paper lies. What all multi-criteria methods have in common is the view that most decisions and decision-making can be improved by decomposing the overall evaluation of alternatives into evaluations on a number of usually conflicting criteria relevant to the problem. Criteria may be defined in quite general terms (e.g. air quality) but should each be associated with a measurable attribute which provides a quantitative or qualitative scale for assessing performance on the underlying criterion. Multi-criteria methods differ primarily according to how they (a) evaluate performances on each attribute, and (b) aggregate evaluations across attributes to arrive at an overall or global evaluation. A recent review of the field can be found in [62].

In many decision problems the evaluation of alternatives is complicated by their performance on at least some attributes not being known with certainty. This uncertainty is intuitively recognizable as a distinct "lack" of complete knowledge or certainty but can derive from many sources and thus assume multiple forms (e.g. the taxonomy in [56]). We use the term "uncertainty" primarily for uncertainty arising when the consequences of an action are unknown because they depend on future events. This is sometimes termed "external uncertainty" [121] because it relates to uncertainty about environmental conditions lying beyond the control of the decision maker. The review therefore excludes many methods aimed primarily at resolving "internal" uncertainty, i.e. uncertainties about decision maker preferences and problem identification (e.g. problem structuring methods [99]), vagueness of information (e.g. rough set theory [68]), and imprecise judgements (e.g. various forms of sensitivity analysis, as well as inverse preference methods such as UTA [115]). We will need to recognize, however, that not all writers differentiate clearly between "internal" and "external" uncertainty, and for this reason the distinction may be blurred in parts of our review.

Facilitating decisions under conditions of uncertainty requires a choice about how this uncertainty is to be modeled. This involves choosing an uncertainty format – a way of representing the possible outcomes that may occur – and a related choice of a decision model with which to represent preferences. Several formats exist for representing uncertainty, and for each of these formats many prescriptive decision models have been developed. This gives a large number of potential approaches for conducting multi-criteria decision analysis (MCDA) with uncertain attribute evaluations. The aim of this paper is to review developments in this area and to organize existing approaches in a coherent way. Because of the sheer number of approaches, any review is unlikely to be exhaustive and we do not attempt this here; rather we have tried first and foremost to do justice to the breadth of the field, and then to provide some depth of coverage by way of additional description and

* Corresponding author.
E-mail address: ian.durbach@uct.ac.za (I.N. Durbach).

0377-2217/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2012.04.038

references to the many models that exist. We have done this by organizing the paper around two main themes: formal models for representing preferences under uncertainty; and the assessment of judgements about uncertain outcomes and preferences associated with these outcomes. Within each of these themes, we organize our review around five uncertainty formats that we found to account for most of the published work on the subject. These (with brief descriptions) are:

Probabilities. A multivariate probability distribution governs the joint realization of performance outcomes across all alternatives and all attributes. Independence conditions may dictate whether the full joint distribution could be constructed from the marginals or whether it would in principle have to be evaluated directly. The use of probabilities and probability distributions, as well as extensions of probabilities known as degrees of belief, is reviewed in Sections 2 and 7.1.

Decision weights. This concept represents an extension of probability. A large body of empirical research suggests that when making decisions people weigh the importance of uncertain outcomes by factors which are typically not linearly related to the associated probabilities; such weighting factors are termed "decision weights". This raises the question of *how* people transform probabilities into decision weights. Research on this question is ongoing but many important results have been obtained. While decision weights can be considered as a generalization of probability, the volume of research on the topic motivates its separate treatment; these findings are reviewed in Sections 3 and 7.1.

Explicit risk measures. Explicit risk measures are attempts to capture the impacts of uncertainties on preferences by means of one or at most a small number of summarized measures. In effect, the assumption is that uncertain attributes can be decomposed into 'value' and 'risk' components. 'Value' components are typically based on expected values or some other central location measure. Specific risk measures are discussed in more detail in Sections 4 and 7.2, but some common examples include variances, ranges, and quantiles.

Fuzzy numbers. Fuzzy set theory is a general theory for the modeling of imprecision but is also often used to treat external uncertainty. The fundamental notion is that imprecision manifests itself as an arbitrariness in establishing precise boundaries for a set of interest, allowing set membership to be considered a matter of degree. Interval assessments of uncertainty may also be treated using fuzzy numbers. Fuzzy decision models are reviewed in Sections 5 and 7.3.

Scenarios. Scenarios are incomplete descriptions of how the future might unfold, with emphasis placed on the development of an internally-consistent chain of causal reasoning that allows the decision maker to gain understanding of the problem at hand and generate unusual insights into possible courses of action. The construction and use of scenarios in decision analysis is reviewed in Sections 6 and 7.4.

In what follows we have attempted to review developments in uncertainty modeling from a general decision-making perspective, without taking the point-of-view of one or other 'school' of multicriteria decision making (e.g. value function, outranking). However, in a few instances where we use mathematical detail to clarify the text, we have used value functions to do so. This is done for compactness only – it is straightforward to translate the value function formulations into formulations appropriate to other schools of MCDA.

In the remainder of the paper, we consider a decision problem consisting of I alternatives denoted by a_i, $i \in \{1, \ldots, I\}$, each evaluated on J criteria denoted by c_j, $j \in \{1, \ldots, J\}$. Let Z_{ij} be the evaluation of a_i in terms of criterion c_j, according to some suitable performance measure. Our concern is with decision making situations in which the values of Z_{ij} for each i are not known with certainty for all j.

2. Decision analysis based on probabilities

In this approach, for each alternative a_i, the Z_{ij} are viewed as random variables with an associated (I-dimensional) multivariate probability distribution function $\mathbf{F}_i$. Let F_{ij} denote the corresponding marginal distribution function for criterion c_j if alternative a_i is selected, and f_{ij} the associated probability density function.

2.1. Multi-attribute utility theory

Expected utility theory (EUT) is almost certainly the most widely-known model for decision making under uncertainty, and for this reason is not reviewed in depth here. Its history can be divided into four periods: (1) early developments around the proof of the expected utility hypothesis for the single-attribute case [136]; (2) reports of systematic descriptive violations of the EUT axioms [5,51]; (3) multi-attribute extensions [82]; and (4) responses to the descriptive violations of EUT [118,4].

The aim of EUT is to produce a 'utility' function U satisfying the 'expected utility hypothesis': that an alternative a will be preferred to another alternative b if and only if the expected utility of a is greater than the expected utility of b. Von Neumann and Morgenstern [136] proved that such a function exists for a single-attribute decision problem provided that a small number of superficially reasonable axioms are satisfied: completeness, transitivity, consistency, continuity, and independence. These axioms form the foundation of subjective EUT. Within a relatively short space of time a number of systematic violations of these axioms, particularly independence, were reported (e.g. the famous paradoxes of Allais [5] and Ellsberg [51]). It is now a well-established fact that subjective expected utility does not adequately describe peoples' choices, in the sense that a substantial proportion of people violate one or more of the axioms at least some of the time.

From the perspective of MCDA, violations of the axioms of expected utility elicit three broad categories of response. The first is that the axiomatic violations are 'not that bad', even in a descriptive sense: that the majority of decision makers are at least in tentative agreement with the axiom of independence [56] and that at least some apparent violations of EUT can be classified as decision making with error rather than axiomatic violation [26]. The second is a pragmatic response: that it is undoubtedly true that expected utility is descriptively invalid, and that violations are frequent and systematic, but that the theory retains a prescriptive value because it allows a simple and coherent framework for constructing preferences in a way that is transparent and easy to explain and understand even to non-technical users (see chapter 4 of [13] for a defense of this argument). The final response views the violations as prescriptively as well as descriptively undesirable, and seeks to extend the expected utility model so that the resulting model is able to accommodate behavior violating the axioms of expected utility, on the basis that the descriptive violations may reveal inadequacies of the expected utility model for representation of human preferences – the subject of Section 3.

In moving from one to several attributes, the aim of (now multi-attribute) utility theory remains to produce a function such that an alternative is preferred to another if and only if its expected utility is greater, but the presence of multiple attributes means that expectations must now be taken with respect to multivariate probability distributions. Practically, this requires (a) the construction, for each criterion c_j, of a marginal utility function u_j satisfying the Von Neumann–Morgenstern axioms, and (b) some way of aggregating the marginal utility functions into a global utility function U such that the expected utility hypothesis is still satisfied. Suitably simple aggregation requires some strong assumptions about the independence of preferences for lotteries on different attributes. For an additive aggregation $U(a) = \sum_{j=1}^{J} w_j u_j(a)$ to be appropriate, preferences for lotteries defined over multiple attributes must be "additively independent", i.e. depend only on the marginal distributions and not on any interactions between attributes. More complex aggregation forms are available if additive independence does not hold (the so-called multiplicative and multi-linear aggregations), but these are only very rarely employed in practice (see [82] for further details).

Because of the complexity of the assessments required by practical applications of MAUT, some pragmatic simplifications have been proposed. Stewart [119] uses a range of simulated problem settings to show that using an additive aggregation when preferences actually follow a multiplicative model may often only have minor impacts on results, but clearly this conclusion cannot be generalized to situations in which strong interactions occur. Von Winterfeldt and Edwards [137, ch. 10] indicate that there may be little difference in using marginal value functions instead of utility functions, which frees the decision maker from specifying preferences between lotteries. Some supporting evidence for this claim has recently been provided in the context of non-expected utility [2].

2.2. Pairwise comparisons of probability distributions

In some instances a pairwise comparison of the associated probability distributions is sufficient to confirm that one alternative is preferred to another (in the sense of maximizing expected utility) provided that certain constraints on the underlying utility function are satisfied. In the case of single-attribute decision problems, three 'stochastic dominance' relations [70,148] are defined as:

$$F_a >_1 F_b \iff H_1(x) = F_a(x) - F_b(x) \leqslant 0, \ \forall x \in [0, \infty) \tag{1}$$

$$F_a >_2 F_b \iff H_2(x) = \int_0^x H_1(y)dy \leqslant 0, \ \forall x \in [0, \infty) \tag{2}$$

$$F_a >_3 F_b \iff H_3(x) = \int_0^x H_2(y)dy \leqslant 0, \ \forall x \in [0, \infty) \tag{3}$$

where F_a is the cumulative single-dimensional distribution function associated with Z_a and the relation $>_i$ refers to first-, second-, and third-degree stochastic dominance for $i = (1, 2, 3)$, respectively. Under the assumption of the expected utility hypothesis, certain preferences can be inferred from first, second and third degree stochastic dominance as follows [11]:

- If $F_a >_1 F_b$ then $a \succ b$ for any increasing utility function;
- If $F_a >_2 F_b$ then $a \succ b$ for any concave and increasing utility function;
- If $F_a >_3 F_b$ then $a \succ b$ for any decreasingly risk averse, concave and increasing utility function.

It should be noted that these results require increasingly strong assumptions, but that the less restrictive lower-degree stochastic dominance conditions are less likely to be observed in practice for any pair of alternatives. Similar conditions have been provided for convex utility functions [61].

In a multi-criteria framework, stochastic dominance conditions can be checked for each individual criterion, using the marginal distributions. Huang et al. [71] have shown that a necessary condition for multi-attribute stochastic dominance is stochastic dominance on each individual criterion. However, the conflicting nature of multi-criteria problems means that these simultaneous conditions are unlikely to occur often.

Stochastic dominance relations have been utilized particularly in the context of outranking methods. Zaras and Martel [161,159] use a simple weighted aggregation of indicator variables $\zeta_j(a_i, a_k)$ which equal 1 if a_i stochastically dominates a_k on criterion c_j and are otherwise zero. This results in a concordance index as for Electre I. Martel et al. [96] and Azondékon and Martel [8] use a more nuanced approach by defining the local preference index $\zeta_j(a_i, a_k)$ as a product of three functions each scaled between 0 and 1 that cause $\zeta_j(a_i, a_k)$ to decrease as dominance conditions weaken from first- to third-order stochastic dominance. A similar threshold-based method is provided in [106], while Zhang et al. [162] also define a numeric stochastic dominance degree which can be used to further discriminate between dominance cases. Dominance-based methods have also been extended to make use of other data types, notably fuzzy numbers, and possibilistic and evidentiary evaluations [160,14,22]. We mention these here because they transform the uncertain quantities so that they assume some of the properties of probability distributions and can thus be compared in much the same way. Notably, this allows for the possibility of using several different data types in the same decision problem.

Other pairwise comparisons of probability distributions have been incorporated into stochastic outranking methods. Jacquet-Lagrèze [74] allocates the part of f_{ij} (f_{kj}) where there is a non-zero probability of a_k (a_i) occuring as evidence in support of indifference, and then uses the cumulative distributions to allocate the remaining probability mass as evidence either that alternative a_i is preferred to a_k, or *vice versa*. Aggregation proceeds as for Electre I. Others [41,96,52,91] compare distributions by constructing a matrix $\mathbf{P}^j$ whose entries P^j_{ik} denote the probability that alternative a_i is superior to alternative a_k on criterion c_j, i.e. $\Pr[Z_{ij} \geqslant Z_{kj}]$. The models differ with respect to the subsequent exploitation of the probabilities. Dendrou et al. [41] and Liu et al. [91] both aggregate the P^j_{ik} using a weighted sum over attributes to arrive at a global index for each pairwise comparison P_{ik}. Fan et al. [52] compute joint probabilities associated with each of 2^J possible permutations of binary indicators denoting (attribute-specific) outranking between a pair of alternatives. Each of these is taken as evidence in favor of the 'superiority', 'inferiority', or 'indifference' of a_i relative to a_k, based on a comparison with a user-defined threshold. A further algorithm is required to exploit the results. Martel et al. [96] incorporate more sophisticated outranking concepts such as indifference and preference thresholds, and subsequent aggregation and exploitation proceeds in a similar fashion to Electre III.

In all of these outranking models the distributional aspect of the problem is fully absorbed into the problem at an early stage of the process through the definition of the P^j_{ik}. In contrast, d'Avignon and Vincke [37] use the uncertain attribute evaluations to form a stochastic (or 'distributive') outranking degree indicating the probability of attaining various degrees of outranking, rather than summarizing the stochastic evaluations directly as P^j_{ik}.

2.3. Models simulating from probability distributions

If evaluations can be specified using probability distributions, one option is to use Monte Carlo simulation to generate values from the distributions and to use these simulated values as inputs to a decision model. Applications of this approach have tended not to differentiate strongly between what we have termed internal and external uncertainties, and often relate to uncertainties in value judgements. This latter context is particularly evident in applications to the analytic hierarchy process (AHP). Early research into the modeling of probabilities in the AHP was largely concerned with deriving relationships between the distributional form of the uncertain pairwise judgements and the distributions of the marginal evaluations contained in the 'priority vector', i.e. the eigenvector corresponding to the largest eigenvalue [134]. Subsequent probabilistic AHP models [90,10] have focused on using Monte Carlo simulation to randomly generate pairwise evaluations from the distributions specified by decision makers. These approaches make small embellishments around the same basic approach. The decision maker expresses pairwise comparisons in the usual way, i.e. using the same 1–9 scale as for deterministic AHP, except that these comparisons are allowed to be random variables with associated probability distributions. No restriction is placed on the types of distributions that are appropriate, so that both external and internal uncertainties (such as imprecision) can be modeled. Next, sets of random pairwise judgements are generated using Monte Carlo simulation. For each set of randomly generated evaluation matrices, the priority vector is computed. Repeating this process many times gives a distribution of priorities for each alternative. These simulated distributions can be used to rank the alternatives, in most cases using the mean of the distribution.

The other area of MCDA in which Monte Carlo simulation from probability distributions has been popular is in the family of stochastic multi-criteria acceptability analysis (SMAA) methods. SMAA is an inverse-preference approach that provides information about the types of preferences that would lead to the selection of particular alternatives. The methods thus focus primarily on internal uncertainty (and so are not described in detail here) but also treat external uncertainty about attribute evaluations by simulating from distributions defined over the values of possible preference parameters as well as possible attribute evaluations. SMAA variants differ in terms of the preference model used and thus the type of preference information that is imprecisely known. Variants are available for value function, outranking, reference point, and prospect theory methods (see the review in [124]).

2.4. Models using belief functions

One criticism of subjective probability is that it does not provide an explicit mechanism for dealing with ignorance. The Dempster–Shafer theory of evidence [113] provides such a mechanism by replacing subjective probabilities with 'degrees of belief' that can be used to represent the extent to which a decision maker believes a specific proposition (for example, that a new drug carries "negligible" side effects) to be true. The degrees of belief assigned to a set of collectively exhaustive and mutually exclusive hypotheses (by a so-called 'belief function') are allowed to sum to less than one, with the difference revealing the degree of ignorance. Such ignorance may be due to a lack of data or familiarity with the problem at hand, imprecision in assessment, or the absence of certain stakeholders in a group decision. Rules for combining evidence from difference sources into a single aggregated assessment are provided in [113].

Belief and belief functions have been introduced into a number of multi-criteria models. Boujelben et al. [22] compare belief functions defined over ordinal assessments using a modified version of first stochastic dominance which employs beliefs rather than probabilities; the same authors also propose comparing belief functions to locally-defined ideal and nadir belief functions using specially-defined distance measures [21]. Perhaps the most widely-known belief-based method for MCDA is the evidential reasoning approach [e.g. 154,151], which treats an alternative's assessment on each attribute as a different source of evidence regarding that alternative's overall suitability. In the basic formulation, the same set of discrete evaluation 'grades' $H = \{H_1, H_2, \ldots, H_N\}$ is used for each attribute. The decision maker expresses their degree of belief $\beta_{nij} \in [0, 1]$ that alternative a_i should be assessed to grade H_n on criterion c_j, with any remaining belief, i.e. ignorance, denoted by $\bar{\beta}_{ij}$. The distributed assessments for each attribute are combined using a recursive algorithm into a global assessment β_{ni} representing the belief that alternative a_i achieves grade H_n, and a global assessment of ignorance $\bar{\beta}_i$. The algorithm, described fully in [156], is based on first multiplying marginal beliefs with importance weights on each attribute, i.e. $m_{nij} = w_j\beta_{nij}$. The probability mass that remains unassigned to any individual grades is then partitioned into what is due to non-zero weights on other attributes, i.e. $m_{ij}^{(w)} = 1 - w_j$, and what is due to incomplete assessment, i.e. $m_{ij}^{(a)} = 1 - \sum_n w_j\beta_{nij}$. The recursive algorithm combines evidence across attributes in such a way that a number of plausible axioms around the 'synthesis' of evidence are satisfied (although see [46] for evidence of descriptive violations). The axioms indicate primarily that no (all) belief should be assigned to an overall grade H_n if no (all) lower-level attributes have been assessed to that grade, and that incompleteness is preserved in the aggregate assessments. Central to the aggregation is the following (given for the two-attribute case but readily extended):

1. The product of m_{nij} with any of m_{nik}, $m_{ik}^{(w)}$ or $m_{ik}^{(a)}$ is taken to be evidence in support of the aggregated assessment m_{ni}.
2. The product of $m_{ij}^{(a)}$ with any of $m_{ik}^{(w)}$ or $m_{ik}^{(a)}$ is taken to be evidence in support of the aggregated assessment of incompleteness $\bar{m}_i$.
3. The product $m_{ij}^{(w)}m_{ik}^{(a)}$ is reallocated back to all individual grades (including the incomplete grade) in proportion to their current assessments, i.e. $\beta_{ni} = m_{ni}/\left(m_{ij}^{(w)}m_{ik}^{(w)}\right)$, $\bar{\beta}_i = \bar{m}_i/\left(m_{ij}^{(w)}m_{ik}^{(w)}\right)$.

Extensions to this basic approach allow for grading systems to differ between attributes as well as for quantitative attributes [e.g. 151]. Both require an additional step in which equivalences are established between overall grades and grades used to describe marginal performances on attributes. These can be done using either rule-based or utility-based transformations [154]. Further extensions allow unallocated beliefs to be redistributed to a restricted grade 'interval', i.e. subset of grades [152], and for beliefs to be expressed as intervals [146] or fuzzy linguistic terms [155] rather than crisp values.

The evidential reasoning algorithm is computationally simple to implement and many applications have been reported in the literature (see the review in [151]). Behavioral issues have not received much attention though, leaving several interesting directions for future research. These include finding meaningful behavioral interpretations (and associated elicitation procedures) for the notion of "degrees of belief", for the allocation of products of different sources of unassigned beliefs, and for the use of different normalization constants at each step.

3. Decision analysis based on decision weights

Whether or not simplifications to multi-attribute utility theory lead to worse prescriptive decisions, the results obtained by Allais [5], Ellsberg [51] and later many others indicate that EUT is not descriptively valid, i.e. does not accurately represent observed choices. This has led others to develop models of choice that are able to capture the elements of choice under uncertainty that expected utility cannot. While these are predominantly descriptive theories, their influence has extended into prescriptive modeling in the search for models that are "more responsive to the complexities and limitations of the human mind" [129]. We begin by reviewing the descriptive developments before considering their implications for MCDA.

3.1. Descriptive theories of choice under uncertainty

There is now a fairly good idea of what properties a decision model should possess if it is to be descriptively valid in the context of a single-criterion decision making problem. Starmer [118] presents a comprehensive review of results obtained from a large number of experimental studies using "probability triangles". These are graphical representations of gambles (or lotteries) defined over three outcomes x_1,x_2,x_3 expressed in utility terms with $x_1 < x_2 < x_3$, so that any gamble can be represented on the triangle using the probabilities of each outcome occuring (p_1, $1 - p_1 - p_3$, p_3). Lines of indifference can be drawn on the triangle between gambles a decision maker is indifferent between, and it is the pattern of these indifference curves that are a major focus of empirical research – because different decision models make different predictions about the properties the indifference curves possess. In the case of standard EUT, the indifference curves linking gambles of equal expected utility u would be represented by $(x_1 - x_2)p_1 + (x_3 - x_2)p_3 = u - x_2$, which for different u describe parallel lines in the probability triangle. Three "stylized facts" are drawn from the literature. The first is that indifference curves fan both inwards and outwards in the same probability triangle. This type of behavior can be accommodated by assuming betweenness rather than independence, an approach followed by some weaker (and hence more practically complex) versions of utility theory (implicit EUT [39]; implicit weighted utility theory [33]). Betweenness still implies linear indifference curves though (the converse is also true), which is contrary to the regularly observed aversion to randomization of equally-valued lotteries (the second stylized fact), which implies non-linear indifference curves [25]. Thus betweenness must be relaxed as well – as has been done in several even weaker non-expected utility models (quadratic utility theory [34]; gamble-dependent utility theory [12]).

Thirdly, violations of EUT tend to be more pronounced when certain or near-certain prospects are involved, i.e. on or near the boundaries of the probability triangle, suggesting the use of a non-linear function to weigh the probabilities according to their importance to the decision maker. The regular pattern underweights large probabilities and overweights small probabilities, suggesting an inverted s-shape form for the probability weighting function [e.g. 131]. Models assigning subjective 'decision weights' to the probabilities of different consequences usually do not satisfy betweenness because depending on the form of the weighting function the sum of the decision weights of complementary events may be sub- or super-additive – and hence dominated alternatives may be selected. In a descriptive setting, this is a distinct advantage because it allows for non-linear indifference curves, but its prescriptive value is less clear. The problem is avoided by further extensions to utility theory (rank-dependent expected utility [108]; cumulative prospect theory [131]; configural weight models [16]). These models assign decision weights based on both the probability of an outcome as well as its rank relative to others in magnitude.

We now turn to a consideration of the relevance of these generalizations to normative or prescriptive decision aid, with a discussion of extensions to multi-criteria decisions.

3.2. Relevance to prescriptive decision aid

Traditionally debate about the prescriptive usefulness of non-expected utility theories has tended towards philosophical argument about the normative status of the various violations of expected utility. As pointed out in [138] however, a motivation for non-expected utility theories can be found without entering this debate. A key part of the practical implementation of utility theory is the elicitation of utilities from the decision maker. A common assumption of many elicitation techniques (e.g. the certainty-equivalent and probability-equivalent methods) is that utilities can be inferred from information provided by the decision maker using the tools of EUT. Bleichrodt et al. [19] call this the "classical elicitation assumption", and make the important points that (a) this is a descriptive assumption, dealing with observed behavior, and is thus independent of the 'normative' debate; (b) imposing the assumption on a decision maker who does not follow EUT can lead to biased assessments of utility.

The practical problems involving the prescriptive use of EUT have led to some exciting developments in the integration of prescriptive decision aid and non-expected utility. These can be divided into three broad areas. The first is the development of alternative assessment techniques that can be used to construct utility functions without using expected utility foundations. The second is the empirical question of how much the classical elicitation assumption biases the assessment of utilities, and the consequential effect on decisions. The third is the development of procedures for the decomposition of multi-attribute non-expected utility functions, similar to those formulated by Keeney and Raiffa [82] for multi-attribute (expected) utility theory.

Significant advances have been made in all three of these areas. Wakker and Deneffe [138] propose a utility assessment method that does not depend on the actual probability values, and is thus robust to the kinds of probability transformations that decision makers often use. Their *gamble-tradeoff method* uses two reference outcomes R and r (with $R > r$) and elicits the value X that makes a decision maker indifferent between the gambles[1] (X, p; r) and (x, p; R) for some x, and the value Y for which $(Y, p; r) \sim (y, p; R)$ for some y. It can be shown that this pair of indifferences can be rewritten in the form $u(X) - u(x) = u(Y) - u(y)$ and thus that a standard sequence of indifference assessments will yield a utility function without needing to know the value of the probability p. Wakker and Deneffe suggest that events without known probability can even be used in place of p, e.g. "surgery will succeed" [138]. All that matters is that the decision maker uses p consistently throughout. Extensions [1,18] allow for the assessment of both the non-expected utility function and the probability weighting function. Bleichrodt et al. [19] develop standard correction procedures for the certainty and probability equivalence methods for cases in which time or other resource constraints prevent interactive discussions between decision maker and analyst. Further extensions in [3] allow the full assessment of the prospect theory utility function, i.e. one that is defined over the whole domain of losses and gains.

Important empirical results have also been contributed. Utility functions obtained using assessment procedures that are sensitive to probability weighting exhibit more risk aversion for gains [138,3] and more risk proneness for losses [3] than utility functions obtained with the gamble-tradeoff method, potentially as a result of certainty effects [138]. The corrective procedures in [19] make fairly large adjustments to both certainty and probability equivalence methods, particularly at high probabilities and particularly for the probability equivalence method. Most evidence supports the form of the probability weighting function proposed by prospect theory: an inverted s-shape for the probability weighting function [1,18], and different weighting functions for losses and gains [1]. Linearity at intermediate probabilities (implying closer correspondence with EUT) is supported in one study [18] but refuted in another [1]. The only model able to elicit full prospect theory utility functions [3] finds utility functions that agree with theory (convex for losses and concave for gains) for just over half of the subjects taking part. Perhaps the most intriguing practical results are found by Abdellaoui et al. in [2]. They compare a number of methods for assessing utility functions and find that (a) there are no inconsistencies between assessment methods once probabilities have been transformed according to prospect theory,

[1] A gamble returning an outcome X with probability p and an outcome r with probability $1 - p$ is denoted by (X, p; r).

and (b) there is no difference between the value functions elicited using riskless strength-of-preference information and the utility functions elicited using more complex choices between lotteries.

Multi-attribute applications require some way of decomposing the multi-attribute preference function into its simpler marginal constituents. In EUT, Keeney and Raiffa [82] were able to use various independence conditions to provide a number of different decompositions – in particular the multi-linear, multiplicative, and additive representations. Miyamoto and Wakker [100] show that many of the same decomposition results obtained under EUT are also valid under non-expected utility. The decomposition procedures use the same definition of utility independence used by Keeney and Raiffa, except that the condition is defined over the set of all rank-ordered gambles rather than *all* gambles. Rank ordered gambles are those whose outcomes have been ranked in descending order of preference, i.e. a two-outcome gamble (x, p; y) is rank-ordered if $x \succeq y$, and includes all deterministic outcomes. Utility independence of each attribute from its complement and full mutual utility independence (i.e. between any subset of attributes and its complement) are necessary and sufficient to infer multi-linear and multiplicative aggregations respectively, as in EUT, but the conditions for an additive representation are somewhat different. This is because marginality – the condition that preferences between gambles depends only on the marginal distributions – imposes linearity in the probabilities and thus implies an expected utility model [50]. Independence results have been used to provide multi-attribute representations for (among others) rank-dependent expected utility and prospect theory in [100], and cumulative prospect theory in [158] and [17]. These developments in non-expected utility suggest that in theory multi-attribute modeling using non-expected utility foundations may not look a great deal different from that obtained using expected utility.

4. Decision analysis based on explicit risk attributes

Given the aim of taking external uncertainty about outcomes on an attribute into account, one possible approach is to use some measure of the consequences of this uncertainty as an attribute in its own right. This approach provides a single (or small number of) risk measures indicating how variable or 'risky' performance is. The fundamental notion is that uncertain evaluations can be expressed in terms of 'value' and 'risk' components. In the value function framework, the models may be formulated as

$$U_i^{(\mathrm{risk})} = \sum_{j=1}^{J}\left[w_j u_j(V_{ij}) - \sum_{k=1}^{\mathcal{K}} w_{ij}^{R^{(k)}} R_{ij}^{(k)}\right] \tag{4}$$

where V_{ij} and $R_{ij}^{(k)}$ are measures of the 'value' and 'risk' of Z_{ij} respectively, w_j is a 'swing weight' reflecting the relative importance of a one-unit change in $u_j(\cdot)$, and $w_{ij}^{R^{(k)}}$ is a weight for $R_{ij}^{(k)}$, termed a 'risk weight' and also interpreted as a swing weight. Note that in this general formulation the risk weights may depend on alternatives as well as criteria. The number of risk measures used is denoted by $\mathcal{K}$ but often there will only be a single risk measure per attribute so that the superscript k could be dropped. The measurement of value is relatively uncontroversial and the use of expected values for this purpose, i.e. $V_{ij} = E[Z_{ij}]$, is widely accepted [111]. There is far less agreement on an appropriate measurement of uncertainty, and there remain several conflicting notions about how uncertainty should be defined and modeled [85].

One way of approaching the problem of risk measurement is to attempt to describe what people mean when they say that an event is 'risky'. Weber and Bottom [147] provide a review of empirical research on how the attributes of gambles influence risk and conclude that: risk increases with an increase in range, variance or expected loss of a gamble; risk decreases if a constant positive amount is added to all outcomes of a gamble; risk increases if all outcomes are multiplied by a constant positive number greater than one; and risk increases if a gamble is repeated many times. These are descriptive points typically based upon a single attribute, and need to be interpreted carefully for the purpose of any normative or prescriptive modeling.

Others have considered links between measures of risk and preference models in the context of single-attribute lotteries [e.g. 77], in much the same vein as the stochastic dominance results reviewed in Section 2.2. For example, variances can be used if the utility function is quadratic or if lotteries are normally distributed and the utility function is monotonically increasing; a linear combination of variance and skewness can be used if the utility function is a third-order polynomial (bearing in mind that polynomial approximations to a utility function will typically only be valid over restricted ranges in view of the usually monotonic nature of the utility function). In single-criterion decision-making the use of variance to measure risk has been criticized [85] for its symmetric nature ("gains" are "losses" are equally penalized) as well as its "ineffective" treatment of low-probability events. Unfortunately, for multi-criteria modeling it seems doubtful that decision makers would be able to give meaningful answers to trade-off questions based on anything more complex that the variance [121]. Using different risk measures for different attributes (where the underlying marginal utility functions are of different functional forms) would also seem to be prohibitively difficult. Multi-attribute applications using variances to assess risk can be found in the context of value function [120] and goal programming methods [9]. Kirkwood [83] has shown that evaluating alternatives by $\sum_{j=1}^{J}\left[w_j u_j(E[Z_{ij}]) - w_{ij}^{R}\sigma_{ij}^2\right]$ with risk weights given by $w_{ij}^{R} = (1/2)w_j u_j''(E[Z_{ij}])$ can lead to close approximations of expected utility under the important conditions that the Z_{ij} be normally distributed and the underlying utility functions "do not deviate too much from linear". Other results [47] suggest that under strongly non-linear preferences this model can perform poorly.

In single-attribute decision modeling, criticism around the use of the variance has led to the development of a number of "downside" risk measures which consider only the impact of negative events. These are comprehensively reviewed in [85]. The semivariance $E[(X - E[X])^2|X < E[X]]$ measures the risk associated with obtaining a below-average performance and has been extended to an expected regret measure [40] using an arbitrary threshold t rather than mean performance, i.e. $E[(X - t)^2|X < t]$. Two further measures of risk can be obtained by either defining an *a priori* desired probability level and assessing the associated quantile of performance (often referred to as 'variance-at-risk' in financial applications), or by defining an *a priori* target and assessing the probability of this target not being met. Others have proposed the probability of loss as a measure of risk [55], or a non-linear combination of the probability of losses and gains [92]. These measures have intuitively appealing associations with everyday usage of the term 'risk', and evaluating alternatives by the chance of a resulting catastrophe has a well-established history in risk analysis [e.g. 54]. However, the use of quantiles (and, by extension, probabilities) for single-attribute risk modeling has been criticized (see [85]) for (a) not accounting for extreme losses beyond the specified cut-off, (b) non-convexity, implying that the risk of a portfolio of alternatives may exceed the sum of the risks of its constituents, and (c) discontinuity with respect to the specified probability level. The implications of these criticisms for MCDA have yet to be established, but it seems clear that the use of any more complex risk measures designed in response to these criticisms – in particular, 'conditional variance-at-risk' measuring expected losses conditional on losses exceeding a specified quantile – runs the risk of placing unrealistic demands on the decision maker's ability to assess inputs and interpret outputs. Indeed we have not found any practical applications of MCDA employing anything more complex than quantiles or probabilities of failing to achieve a target.

Multi-criteria applications using lower quantiles to measure risk can be found in [120]; triples consisting of the minimum, median/mode, and maximum are also popular in fuzzy decision analysis [e.g. 87] although the extremes have also been shown to be subject to strong assessment biases (see Section 7.2, also [6]). In practical decision analysis it is common to estimate probability distributions using a small number of (usually three to five) quantiles, this being the basis for well-known elicitation methods like the bisection or interval methods [e.g. 117]. Keefer and Bodily [81] have also shown that the single-attribute expected utility of an alternative can be closely approximated by assessing utilities at each of the 5%, 50%, and 95% quantiles of performance and forming a weighted sum in which the median receives a weight of 0.63 and the 5% and 95% quantiles each receive a weight of 0.185. Note, however, that their approach used quite moderate probability levels. Many seriously risky decisions involve much smaller probabilities which are notoriously difficult to assess and to interpret. In recent simulation [49] and behavioral studies [48] comparing the effects of various uncertainty formats on decision making, the use of a small number of quantiles returned favorable results across a range of decision problems, and in the case of unfacilitated decision making led to better and less cognitively demanding decisions than probability distributions, although these again did not involve very extreme probabilities.

Multi-criteria applications using probabilities of below-target performance to measure risk can be found for goal programming [38] and value function methods [120]. However, the use of a single fixed target means that there is no guaranteed existence of an equivalent utility function formulation for these probability-based risk measures [27]. In order for such an equivalence to exist, the target must be probabilistic. Castagnoli and Calzi [27] show that an alternative formulation of the expected utility model is to assume a decision maker who has only two different utility levels depending on whether an uncertain target is met or not. The 'target-oriented' decision maker assesses probabilities $p(x)$ that the target is achieved given an attribute performance of x, rather than a utility function $u(x)$. Bordley and Kirkwood argue in [20] that in some circumstances this may be a "more intuitively appealing task", and extend the single-attribute results in [27] to show that for each multi-linear (or multiplicative or additive) utility function, there is an equivalent multi-linear (or multiplicative or additive) target-oriented formulation. In fact both the variance-based and probability-based goal programming models can be shown to be special cases of the target-oriented preference model [20].

5. Decision analysis based on fuzzy numbers

Fuzzy decision models seek to model the elements of the decision-making process that are subject to uncertainty using fuzzy sets and numbers. From a philosophical point-of-view, uncertainty in any of these elements is said to arise because of (a) unquantifiable information, (b) incomplete information, (c) nonobtainable information, and (d) partial ignorance [32], although not natural random processes. Practically though, these different uncertainty types are often treated in similar ways. Several authors divide fuzzy multi-attribute decision making into two stages (e.g. [32]). The first stage consists of the assessment and aggregation of attribute evaluations. As the resulting evaluations are usually fuzzy numbers themselves, their ranking is often non-trivial and some kind of ranking procedure is needed in the second stage. There are a large number of ranking methods – certainly more than 50, based on reviews in [32,139,140] and our own more recent searches – and these can almost always be chosen independently of the choice of aggregation model in the first stage. We therefore do not attempt to review fuzzy ranking methods here.

Many if not all decision models can be fuzzified because of the existence of fuzzy versions of nearly all operations that are likely to be employed by a decision model (e.g. addition, multiplication, finding a minimum or maximum). Since most operations can be fuzzified in different ways, fuzzy models have proliferated by employing various permutations of fuzzy operations, giving the field a somewhat chaotic appearance. Four decision models that have received the most attention from fuzzy researchers are focused on: weighted additive models, models based on ideas related to the analytic hierarchy process, aspiration-based models using comparisons to constructed ideal and anti-ideal solutions, and models based on rough set theory.

5.1. Models using weighted additive sums

The standard fuzzy additive value function model uses fuzzy numbers to represent attribute values and/or attribute weights before aggregating performance using fuzzy analogs of multiplication and addition. A number of such methods are reviewed in [32], and more recent applications can be found in [30,36]. Models differ mainly with respect to the way in which the fuzzy numbers formed using the weighted additive aggregation are exploited to arrive at a ranking. The case of solving for unknown attribute weights is also fairly popular [e.g. 145], but falls outside the scope of the current review.

5.2. Models using the analytic hierarchy process

The AHP has received substantial attention from fuzzy practitioners. On the face of it, the AHP seems an obvious candidate for fuzzification because of the qualitative nature of the pairwise comparisons made. Although these judgements are traditionally quantified on a 1–9 scale, the points along the scale are given linguistic meanings and comparisons will often be made predominantly with these linguistic labels in mind. It is therefore easy to imagine that a decision maker might be internally uncertain about some of these judgements; external uncertainties appear to be treated in the same way. Since the 1996 review in [126], we have found more than 30 papers involving the application of fuzzy set theory to the AHP, and there are doubtless many more. The methodological papers differ predominantly in how the vector of marginal preference information is computed from the matrix of judgemental ratios assessed by the decision maker.

It is well-known that because the judgements expressed by the decision maker in AHP may not be consistent, there are a variety of ways to estimate the marginal information. The original approach [110] uses the principal eigenvector of the matrix of assessed judgements, but two further estimation approaches are also employed, using least squares [76] or logarithmic least squares optimization [86]. All three estimation approaches have been modified to make use of fuzzy input data: eigenvalue methods in [126,141], least squares in [153], and log least squares in [143]. Several other methods have also been proposed. Marginal evaluations are based on fuzzy geometric means in [23]. Deng [42] uses fuzzy arithmetic means, and forms crisp evaluations by computing a weighted average of the left and right points of specified α-cuts of each fuzzy mean. These are then evaluated against positive and negative ideal solutions. Mikhailov et al. [97,98] uses a preference programming formulation to maximize the consistency of judgements and the possibility of those judgements. Leung and Cao [89] define a different fuzzy version of consistency based on specifying a tolerance level for inconsistent ratio judgements, and maximize membership values subject to constraints on this consistency. Yu et al. [157] use a multi-objective linear programming approach to maximize membership values and minimize inconsistency, requiring an additional trade-off parameter. Chang's 'extent analysis' [29] uses fuzzy arithmetic means to represent marginal evaluations and then performs pairwise comparisons by computing the possibility that one fuzzy mean is greater than another. This approach is currently widely cited, with some 20 papers published since 2002. However, Wang et al. [144] have recently convincingly argued that the weights obtained using extent analysis model do not represent the relative importance of criteria or alternatives because they confuse evidence of strength with the measurement of that strength.

5.3. Models using comparisons to ideal solutions

Another class of models that have been relatively popular targets for fuzzification are those that evaluate alternatives by comparing them to an ideal and/or anti-ideal solution and choose the alternative that is in some sense closest to the ideal and farthest from the anti-ideal. Most of these models make use of the TOPSIS method [72]. This begins by defining two hypothetical alternatives: an ideal solution consisting of the maximum weighted evaluations across all alternatives on each attribute, and an anti-ideal solution consisting of the minimum evaluations. Euclidean distances between each alternative and the ideal and anti-ideal solutions are computed and alternatives are evaluated based on its distance to the ideal solution expressed as a proportion of the sum of the two distances. All that is required in order to use fuzzy input values are computations of fuzzy maxima (for the ideal solution), minima (for the anti-ideal solution), and distances. All of these are standard fuzzy operations.

Differences between the fuzzy TOPSIS methods relate primarily to when in the decision process the fuzzy information is condensed into crisp evaluations. Tsaur et al. [127] do this right at the beginning, representing fuzzy evaluations by their centroids and calculating the crisp distances accordingly. Chu and Lin [35] use fuzzy multiplication to obtain fuzzy weighted ratings but then convert these into crisp values using the method of mean removals [80]. Chen [31] defines a crisp 'Euclidean' distance between two triangular fuzzy numbers, following which the deterministic TOPSIS method can be applied. Ashtiani et al. [7] use much the same approach, but apply it to interval-valued fuzzy numbers, i.e. a fuzzy number that is defined by two membership functions, one 'upper' and one 'lower'. Separate computations are carried out using the upper and lower membership functions and at the final step a simple average is taken of the relative idealities using the two membership functions. Triantaphyllou and Lin [126] use fuzzy arithmetic operations at each step, so that the result is a fuzzy relative ideality for each alternative that must be ranked using one of the fuzzy ranking methods. Their method preserves fuzziness right up to the ranking stage, but Wang and Elhag [142] find that this results in the supports of the fuzzy relative idealities being overexaggerated. Their method uses a series of α-cuts, calculating the relative ideality of each alternative at each α-cut. The relative idealities are each interval numbers whose lower and upper bounds can be found using a simple fractional programming model and then collected across α-cuts.

5.4. Models integrating fuzzy and rough sets

Rough set theory is a rule-based approach primarily used to treat internal uncertainty arising from information granulation or 'vagueness' rather than external uncertainty, and for that reason we provide only a highly condensed summary here (see the review in [68]). Its key observation is that some alternatives may be equivalent from the point of view of one subset of attributes, i.e. if they share the same attribute values, *even though they differ on another subset of attributes*. Every subset of attributes $\mathcal{J}$ thus partitions the set of alternatives into groups that are equivalent (or "indiscernible') from the point of view of $\mathcal{J}$, called equivalence classes. Any subset of alternatives $\mathcal{I}$ can then be approximated using knowledge contained in $\mathcal{J}$. This approximation takes the form of two crisp sets: a $\mathcal{J}$-lower approximation containing all alternatives in $\mathcal{I}$ that do not share an equivalence class with any alternatives outside $\mathcal{I}$; and a $\mathcal{J}$-upper approximation consisting of any alternatives which share an equivalence class with at least one member of $\mathcal{I}$. Many of the developments in rough sets center around the approximation of equivalence classes defined using holistic or global judgement (i.e. on a 'decision' attribute) using equivalence classes defined using multiple 'condition' attributes. The theory allows the extraction of if-then decision rules which can be used to generate understanding as well as classify new alternatives to classes of the decision attribute using condition attributes.

A rich theory and practice has arisen from rough set theory. Replacing indiscernibility with a less-restrictive 'similarity' relation allows imprecision in assessments [116]. Extensions to ordinal evaluations (critical for multi-criteria sorting problems) are based on approximating upward and downward unions of classes, rather than the classes themselves [64]. Further extensions to multi-criteria ranking and choice problems replace indiscernibility relations with 'dominance' relations expressing outranking relationships defined over pairs of alternatives on subsets of attributes [66]. In [69] this approach is applied to a single-criterion decision problem under uncertainty by considering discrete states underlying stochastic attribute evaluations as criteria in their own right. Missing data can be readily incorporated [65]. Our view is that in spite of these extensions, rough set theory remains primarily a supervised learning approach and that, for decision analysis, the requirement that at least *some* global judgements be given may limit its applicability in large-scale problems where it is precisely because of the difficulty of global judgements that decision support is often sought.

Fuzzy and rough sets can be integrated by fuzzifying the indiscernibility, similarity, or dominance relations, i.e. allowing these to be considered a matter of degree rather than binary [45,67]. This requires a fuzzy extension to the definition of lower and upper approximations using fuzzy connectives, which in turn means that the $\mathcal{J}$-lower and $\mathcal{J}$-upper approximations are themselves fuzzy sets. Greco et al. [63] point out that because of the existence of several fuzzy connectives there is no generally accepted method for doing this, and use this fact to motivate a more conservative approach that considers only ordinal membership information and no connectives, but does require the potentially tricky assessment of functions relating the credibility of membership in condition attributes with the credibility of membership in decision attributes.

6. Decision analysis based on scenarios

As we move into the realm of strategic decision problems, uncertainties become increasingly complex and interrelated, with the result that the elicitation of measures such as probabilities, belief functions or fuzzy membership functions become operationally difficult for decision makers to comprehend, and virtually impossible to validate. In such contexts it becomes useful, especially for external uncertainties, to construct a number of narratives which describe possible ways in which the future might unfold. Each of these possible futures is then conventionally termed a 'scenario'. The most fundamental requirement of such scenarios is that they be internally consistent – that is, that they present coherent descriptions of future worlds without contradictory elements.

The use of scenarios for strategic planning was developed as the more-or-less qualitative technique of scenario planning [e.g. 132], which also gave the term 'scenario' its more specific meaning. In the scenario planning view, the constructed set of scenarios are useful in their own right in getting a group of decision makers to express their views on uncertainty and agree on a small number of progressions to future states of the world. It may be possible to evaluate these scenarios in terms of their desirability (introducing a multi-criteria aspect; see [73]), even though the decision makers have no control over which scenario occurs. More commonly though, the constructed scenarios are used to evaluate and develop alternative strategies. It is at this stage that multi-criteria modeling seems applicable, but many advocates of scenario planning prefer to avoid formal quantitative modeling [e.g. 132,112] and use informed but informal judgement – examples can be found in [149,24].

Nevertheless efforts have been made to integrate decision analysis with the use of scenarios, as discussed for example in Chapter 16 of [60]. The main objective of a scenario-based MCDA model is to evaluate and compare the performances of alternatives in each of N scenarios – given a decision problem, the approach considers that problem separately in each scenario before (possibly but not necessarily explicitly) combining this information to arrive at a final decision. The general approach can be represented using a value tree in which scenarios are placed in the second level of the hierarchy as parents to N structurally similar "within scenario" value trees. Although from a mathematical perspective scenarios can also be included at lower levels of the objectives hierarchy, this may be considered contrary to the philosophy of scenario planning.

If a value function framework is adopted, and if the metacriteria are preferentially independent, then an additive aggregation giving a scenario-based utility for alternative a_i would be defined, following [121], by

$$U_i^{(\text{scen})} = \sum_{k=1}^{N}\sum_{j=1}^{J} w_{jk} u_{jk}(z_{ij}^{(k)}) \tag{5}$$

where $z_{ij}^{(k)}$ is the evaluation of alternative a_i on criterion c_j in scenario s_k, and w_{jk} and u_{jk} are respectively the weight and marginal utility function associated with criterion c_j under the assumption of scenario s_k. If preferences can be assumed constant over scenarios (see Section 7.4 for a discussion of this) then the evaluation simplifies to

$$U_i^{(\text{scen})} = \sum_{k=1}^{N}\left[\omega_k \sum_{j=1}^{J} w_j u_j(z_{ij}^{(k)})\right] \tag{6}$$

where ω_k is the weight associated with scenario s_k.

One of the most important issues in integrating the use of scenarios with multi-criteria analysis is how (or even whether) to compare and aggregate results from different scenarios. Some applications emphasize results within scenarios and prefer not to aggregate evaluations over scenarios to arrive at a final global evaluation [60,109]. While this is in line with the philosophy of scenario planning, which has strong 'robustness' views [132], it is not evident that unaided intuitive aggregation across scenarios will generate the most goal-directed solutions. It remains unclear exactly how more formal aggregation is to be done. Some applications have aggregated results over scenarios using some form of relative likelihood [84], but it has been argued [121] that aggregation should not use scenario "probabilities" (because the set of scenarios does not constitute a complete probability space), nor "likelihoods" (because scenarios are incomplete descriptions, they cannot in general be expected to represent the same dimensions in probability space). Rather, the aggregation should use relative ("swing") weights on performance in different scenarios. This is theoretically permissible because performance within each scenario constitutes an interval preference scale, although "it may be difficult to elicit appropriate values for the scenario weights" [121]. The use of swing weights (as well as probabilities) has been criticized [109], on the argument that it is contrary to the spirit of scenario planning and thus runs the risk of diluting results and alienating users. Recent applications [101,109,133] have thus focused on assessing the stability of performances over scenarios by directly presenting the decision maker with information about the range of performances across scenarios. These are either based on absolute performance [101] – augmented by an aggregation which maximizes the worst performance across scenarios – or regret [109,133] – in which case a robust alternative is defined as one with relatively small regret compared to the alternatives across a wide range of plausible futures [88] – possibly aggregated over scenarios using a second min–max regret computation [133]. In current work [122] we have, however, demonstrated that in some cases at least such robustness-based arguments may be misleading.

7. Assessment, interpretation and framing effects

Decisions under uncertainty require the decision maker to make various judgements about uncertain quantities: which outcomes are believed possible, which are more likely to occur, and so on. In this final part of our review, we wish to focus on the assessment and interpretation of the inputs and outputs of the uncertainty models described in previous sections, in order to identify both what is known and what is, in our opinion, as yet unresolved. Many of these findings are drawn from behavioral decision research, a field with strong but sometimes neglected ties to MCDA (see [102] for a summary).

7.1. Probabilities and related quantities

Behavioral research into the assessment and use of probabilities provides a good example of how descriptive results can improve prescriptive practice, and provides something of a benchmark against which to judge other approaches to modeling uncertainty. Firstly, a detailed and well-accepted body of knowledge [130,59] documents how unassisted probability judgments can suffer from systematic error. The "heuristics and biases" model proposes that people make use of a number of simplifying heuristics when assessing probabilities. Well-known heuristics include anchoring and adjustment, availability, representativeness, and affect. Systematic errors can arise from the use of any of these heuristics. As a consequence today very few, if any, analysts would consider it acceptable practice to simply ask a decision maker to state a subjective probability and trust that estimate as "good", in the sense [e.g. 57] of being (a) internally consistent or coherent, i.e. obeying certain logical conditions like the laws of probability, (b) externally consistent or well-calibrated, i.e. against any available data, and (c) self-consistent or reliable, i.e. over repeated tests. Practical methods for eliciting probability distributions – for example, probability wheels and related visual aids, or the use of additional questions as consistency checks, etc. – have been a cornerstone of MCDA practice for several decades [e.g. 117,137,13].

Furthermore new findings on unassisted probability judgments provide an ongoing basis for evaluating assessment procedures in MCDA, and adapting these where necessary. A good example is the gamble-tradeoff method of utility elicitation [138]. This method, which does not require the exact specification of probabilities, is a direct response to findings on the systematic over- and under-weighting of small and large probabilities observed in descriptive studies [131]. A summary of other evidence [114,78] continues to suggest that good probability assessments can be obtained under current best practice in MCDA. We consider the three related elements described above – knowledge of what influences judgments and an at least tentative model of how these judgments are formed; the development of rigorous and generally-accepted assessment procedures; and responsiveness to new descriptive findings – to be central to a scientific approach to the assessment of uncertainty (and indeed other quantities) in MCDA. Judged by these criteria, many opportunities for filling in the gaps in our knowledge of other uncertainty approaches present themselves.

7.2. Explicit risk attributes

Research on the assessment of explicit risk measures has focused on quantiles and ranges obtained from quantiles, and variances. Many of the findings relating to probability distributions apply to the assessment of quantiles as well. Experiments have found that people can often estimate quantiles reasonably well [57], although there is a clear tendency for people to overestimate their ability to predict an uncertain quantity so that their interquantile intervals tend to be too narrow [130]. There is substantial uncertainty about which quantiles people are most accurate at assessing. Bisection methods can be used to obtain medians and upper and lower quartiles, with good results reported in [104]; other studies find that overconfidence is reduced if the 33% and 67% quantiles are used [58]. There is more agreement on the relative difficulty of accurately assessing the tails of the distribution [6], casting some doubt on the appropriateness of using extreme quantiles. As before, practice, training and feedback can all assist to improve the accuracy of assessments [57]. Variances on the other hand are usually found to be difficult to interpret and assess numerically [57], suggesting that it may be preferable to approximate these, where required, from the quantiles using the results in [81].

7.3. Fuzzy numbers

The assessment of membership functions has received much less attention than it deserves given their obvious centrality to fuzzy modeling and the subtlety of the concept of membership. This is an ongoing criticism of fuzzy decision aid [13] which also applies to other parameters of fuzzy decision and ranking methods that are often less-than-clearly defined. Although the broad picture has not changed greatly [15], the problem does seem to be attracting somewhat more attention in recent years. Verkuilen [135] identifies three methods for assessing membership functions, which are slightly elaborated upon in [15]. The first is direct assignment, in which decision makers offer direct judgements on numerical membership values. This continues to be the most popular method in practice [135] although it raises obvious (and unanswered) empirical questions about whether these judgements can be made reliably and without excessive difficulty. It has also been shown to return membership functions with wider support [28], and estimates that are biased towards the end-points of the membership function [125], relative to other assessment methods.

The second method, closely related to the first, is to assess membership functions by using some theoretically-informed transformation to map a suitable numerical variable onto membership values [e.g. 75]. This merely transfers subjectivity around the assessment of membership to subjectivity around the choice of transformation and numerical input. Trapezoidal and triangular membership functions are widely used [e.g. 87] but often with little meaningful justification. Marchant [95] has recently identified a set of no less than 13 axiomatic requirements which must be satisfied in order for the trapezoidal representation to be valid. Many of these would be complex to assess and it remains an open question whether they are satisfied by decision makers, either descriptively or prescriptively. Related objections to the generic use of interval numbers (uniform distributions) and triangular distributions have also been made in [107,57], essentially claiming that these are in most cases overly simplistic.

The final method is indirect assignment, in which decision makers are asked for some other information (typically cognitively easier to assess, like pairwise comparisons) which is then used to generate membership values via some additional operations. Marchant [93,94] has again provided conditions under which it is possible to construct membership functions using either pairwise comparisons of membership differences or assessments of the ratios of membership values. Some authors have argued against ratio-scaled assessments [28], but the general challenge for future research is to assess whether these axiomatic requirements, many of which are quite complex, are empirically fulfilled. That is, do these axioms reflect what decision makers mean (or 'should' mean, a more difficult prospect) when they assess and compare set membership? The little evidence that exists on this question gives little cause for optimism, with Desimpelaere [43] and Desimpelaere and Marchant [44] showing empirical violations, albeit with small sample sizes, of several axioms required in order to use max and min operators to represent the union and intersection of fuzzy sets, respectively.

7.4. Scenarios

Most scenario planning texts deal explicitly with the question of scenario construction, so that desirable features of scenarios [150] as well as techniques for their construction [132,112] are well-established and fairly widely accepted, and are therefore not considered here. Instead, we focus on the generally quantitative construction of performance measures within each scenario that must be assessed as part of scenario-based MCDA. Two general procedures are given in [121]. The first considers combinations of alternatives and scenarios as IN distinct outcomes or 'meta-alternatives' to be evaluated in terms of the J attributes. A marginal preference model (whether this uses a value function, outranking, goal programming, or other approach) is defined across all IN outcomes for each attribute. The result is an $I \times N$ table giving the aggregated (over attributes) performance of alternative a_i under scenario s_k. These evaluations may then be aggregated over scenarios although as discussed in Section 6 there is currently no consensus on the best way to do this. The second procedure considers combinations of scenarios and attributes as NJ distinct 'meta-attributes', and evaluates the I alternatives in terms of each of these meta-attributes. This means that a marginal preference model is constructed for each of the NJ meta-attributes, following which performance would be aggregated over all meta-attributes (possibly first within each scenario and then over scenarios, if this is desired). Weights for each scenario-attribute combination can be assessed directly (although this may be impractical for large numbers of attributes or scenarios) or by first establishing the weight of each criterion c_j under the assumption of a common scenario s_k, denoted $w_{j|k}$, standardizing these to sum to one within each scenario, and then assessing the weight ω_k associated with each scenario s_k. The joint weights are given by the product $\omega_k w_{j|k}$.

It is worth noting that it is only in the second assessment procedure that preference information is allowed to vary between scenarios. One application of scenario-based MCDA [101] found that progress was only possible once importance weights were allowed to vary across scenarios – and hence the second model evaluating alternatives over scenario-attribute pairs was used. Preference information is also allowed to vary across scenarios in [84] but not in [60]. It remains an open question how often the detailed qualitative information gathered during the construction of the scenarios might cause scenario-dependent preferences, or at least an awareness of those preferences. Another issue is what kinds of heuristics and biases may affect these assessment procedures. There appears to be no research directly focusing on this area, and the key question of whether the nature of the underlying scenarios might bias evaluations in some way remains open. Evidence suggests that people focus on optimistic scenarios and ignore pessimistic scenarios when predicting their own performance, and also rate pessimistic scenarios as less likely to occur [105], which might conceivably result in assessments in favorable scenarios being more reliable than those in unfavorable scenarios, or bias relative trade-offs between performance levels in different scenarios if these are aggregated probabilistically. Currently though, it is simply not known how and even which qualitative characteristics of scenarios, if any, influence the quantitative assessment of performance within scenarios.

8. Concluding comments

Some significant achievements in the modeling of uncertain attribute evaluations for MCDA problems have been made in the past decade. From the perspective of encouraging greater application of MCDA in strategic decisions involving uncertainty, two developments which hold particular promise are practical procedures for the implementation of non-expected utility and the scenario-based MCDA approaches. Non-expected utility approaches have long been recognized as capturing a broader range of empirically observed behavior than EUT, but their adoption as prescriptive tools has been slow at least partly because of the complexity of their implementation. Scenario planning is a popular and comprehensive decision support tool for considering future uncertainties, but its mechanisms for evaluating alternatives are limited and can appear *ad hoc*. The opportunity exists for MCDA to provide scenario planning with a structured, systematic approach for evaluating alternatives, and in doing so to use scenario planning's popularity to increase its own. The few applications of scenario-based MCDA to date suggest that this integration has met with some success.

These two methods also exemplify reasonable compromise positions on the spectrum between methods that are theoretically rich but heavily parameterized and practically complex, and those that are transparent and easily understood but may not conform to prescriptive principles of rationality. The search for a balance between these conflicting goals continues to define and indeed polarize much of MCDA research. Behavioral decision theory, the study of human judgement and decision, has become more influential as MCDA tools have adapted in an attempt to reduce the effects of known response biases in their inputs [102]. This has mostly taken the form of increasing attention paid to problem structuring, to the extent that a list of conditions minimizing response biases would overlap considerably with a list of current best practice for problem structuring in a prescriptive decision analysis. Yet our impression in conducting this review is that many models treat assessment and interpretation issues only superficially, and tend to focus on mathematical and algorithmic detail rather than transparency to the decision maker. Working towards a research culture that insists that in order to be useful (and indeed publishable) a decision model must possess transparency, rigorous assessment procedures and a clear, unambiguous behavioral interpretation for all quantities and parameters employed should be a priority for the field. In the meantime, assembling these crucial missing elements of several existing models offers a number of worthy avenues for research.

The effect of the observed growth in available decision models on the use of MCDA in practice is not yet clear, but it appears that the increased number of decision models has until now not been matched by a proportional increase in the number of real-world applications of MCDA reported in the literature. The number of available models is both a opportunity – because all else being equal it allows a greater variety of decision makers and decision problems to be accommodated – and a threat – because without careful explanation the whole process may be viewed as confusing, disjointed, and seemingly arbitrary. Similar issues are faced by different schools of preference modeling in MCDA (e.g. value function and outranking), but whereas these have received a great deal of attention [e.g. 123,128] relatively little work has been done on organizing and reconciling models for uncertain decision making. Decision makers would perhaps best be served by future research increasing the depth rather than the breadth of the field, by exploring the relationships between existing uncertain decision models and the extent to and conditions under which they differ, and organizing these models into a framework from which practical guidelines can be developed for choosing between them in a transparent way.

References

[1] M. Abdellaoui, Parameter-free elicitation of utility and probability weighting functions, Management Science (2000) 1497–1512.
[2] M. Abdellaoui, C. Barrios, P. Wakker, Reconciling introspective utility with revealed preference: Experimental arguments based on prospect theory, Journal of Econometrics 138 (1) (2007) 356–378.
[3] M. Abdellaoui, H. Bleichrodt, C. Paraschiv, Loss aversion under prospect theory: a parameter-free measurement, Management Science 53 (10) (2007) 1659–1674.
[4] M. Abdellaoui, J. Hey, Advances in Decision Making Under Risk and Uncertainty, Springer Verlag, 2008.
[5] M. Allais, Le comportement de l'homme rationnel devant le risque: Critique des postulats et axiomes de l'école Americaine, Econometrica 21 (1953) 503–546.
[6] M. Alpert and H. Raiffa. A progress report on the training of probability assessors. In Judgment under uncertainty: Heuristics and biases [79] pp. 306–334.
[7] B. Ashtiani, F. Haghighirad, A. Makui, G. ali Montazer, Extension of fuzzy TOPSIS method based on interval-valued fuzzy sets, Applied Soft Computing 9 (2009) 457–461.
[8] S. Azondékon, J. Martel, Value of additional information in multicriterion analysis under uncertainty, European Journal of Operational Research 117 (1999) 45–62.
[9] E. Ballestero, Stochastic goal programming: a mean-variance approach, European Journal of Operational Research 131 (2001) 476–481.
[10] R. Banuelas, J. Antony, Application of stochastic analytic hierarchy process within a domestic appliance manufacturer, Journal of the Operational Research Society 58 (1) (2007) 29.
[11] V. Bawa, Optimal rules for ordering uncertain prospects, Journal of Financial Economics 2 (1) (1975) 95–121.
[12] J. Becker, R. Sarin, Gamble dependent utility, Management Science 33 (1987) 1367–1382.
[13] V. Belton, T. Stewart, Multiple Criteria Decision Analysis: An Integrated Approach, Kluwer Academic Publishers, Boston, 2002.
[14] S. Ben Amor, K. Jabeur, J. Martel, Multiple criteria aggregation procedure for mixed evaluations, European Journal of Operational Research 181 (3) (2007) 1506–1515.
[15] T. Bilgiç, İ. Türkşen, Measurement and elicitation of membership functions, in: W. Pedrycz, A. Skowron, V. Kreinovich (Eds.), Handbook of Granular Computing, Wiley Online Library, 2008.
[16] M. Birnbaum, New paradoxes of risky decision making, Psychological Review 115 (2) (2008) 463.
[17] H. Bleichrodt, J. Miyamoto, A characterization of quality-adjusted life-years under cumulative prospect theory, Mathematics of Operations Research (2003) 181–193.
[18] H. Bleichrodt, J. Pinto, A parameter-free elicitation of the probability weighting function in medical decision analysis, Management Science (2000) 1485–1496.
[19] H. Bleichrodt, J. Pinto, P. Wakker, Making descriptive use of prospect theory to improve the prescriptive use of expected utility, Management Science (2001) 1498–1514.
[20] R. Bordley, C. Kirkwood, Multiattribute preference analysis with performance targets, Operations Research 52 (6) (2004) 823.
[21] M. Boujelben, Y. De Smet, A. Frikha, H. Chabchoub, A ranking model in uncertain, imprecise and multi-experts contexts: the application of evidence theory, International Journal of Approximate Reasoning 52 (8) (2011) 1171–1194.
[22] M. Boujelben, Y. Smet, A. Frikha, H. Chabchoub, Building a binary outranking relation in uncertain, imprecise and multi-experts contexts: the application

of evidence theory, International Journal of Approximate Reasoning 50 (8) (2009) 1259–1278.
[23] J. Buckley, Fuzzy hierarchical analysis, Fuzzy Sets and Systems 17 (3) (1985) 233–247.
[24] G. Cairns, G. Wright, R. Bradfield, K. van der Heijden, G. Burt, Exploring e-government futures through the application of scenario planning, Technological Forecasting & Social Change 71 (3) (2004) 217–238.
[25] C. Camerer, T. Ho, Violations of the betweenness axiom and nonlinearity in probability, Journal of Risk and Uncertainty 8 (2) (1994) 167–196.
[26] E. Carbone, J. Hey, Estimation of expected utility and non-expected utility preference functionals using complete ranking data, in: Munier and Machina [103], pp. 119–140.
[27] E. Castagnoli, M. Calzi, Expected utility without utility, Theory and Decision 41 (3) (1996) 281–301.
[28] J. Chameau, J. Santamarina, Membership functions I: comparing methods of measurement, International Journal of Approximate Reasoning 1 (3) (1987) 287–301.
[29] D. Chang, Applications of the extent analysis method on fuzzy AHP, European Journal of Operational Research 95 (3) (1996) 649–655.
[30] T.-H. Chang, T.-C. Wang, Using the fuzzy multi-criteria decision making approach for measuring the possibility of successful knowledge management, Information Sciences 179 (2009) 355–370.
[31] C. Chen, Extensions of the TOPSIS for group decision-making under fuzzy environment, Fuzzy Sets and Systems 114 (1) (2000) 1–9.
[32] S.-J. Chen, C.-L. Hwang, Fuzzy Multiple Attribute Decision Making: Methods and Applications, Springer-Verlag, Berlin, 1992.
[33] S. Chew, Axiomatic utility theories with the betweenness property, Annals of Operations Research 19 (1) (1989) 273–298.
[34] S. Chew, L. Epstein, U. Segal, Mixture Symmetry and Quadratic Utility, Econometrica 59 (1) (1991) 139–163.
[35] T. Chu, Y. Lin, A fuzzy TOPSIS method for robot selection, The International Journal of Advanced Manufacturing Technology 21 (4) (2003) 284–290.
[36] T.-C. Chu, Y. Lin, An extension to fuzzy MCDM, Computers and Mathematics with Applications 57 (2009) 445–454.
[37] G. d'Avignon, P. Vincke, An outranking method under uncertainty, European Journal of Operational Research 36 (1988) 311–321.
[38] P. De, D. Acharya, K. Sahu, A chance-constrained goal programming model for capital budgeting, Journal of the Operational Research Society 33 (1982) 635–638.
[39] E. Dekel, An axiomatic characterization of preferences under uncertainty: weakening the independence axiom, Journal of Economic Theory 40 (2) (1986) 304–318.
[40] R. Dembo, D. Rosen, The practice of portfolio replication. A practical overview of forward and inverse problems, Annals of Operations Research 85 (1999) 267–284.
[41] B. Dendrou, S. Dendrou, E. Houstis, Multiobjective decision analysis for engineering systems, Computers and Operations Research 7 (1980) 301–312.
[42] H. Deng, Multicriteria analysis with fuzzy pairwise comparison, International Journal of Approximate Reasoning 21 (1999) 215–231.
[43] C. Desimpelaere, Empirical Validation of the Measurement-Theoretic Foundations of Fuzzy Set Theory, PhD thesis, Ghent University, 2005.
[44] C. Desimpelaere, T. Marchant, An empirical test of some measurement-theoretic axioms for fuzzy sets, Fuzzy sets and systems 158 (12) (2007) 1348–1359.
[45] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, International Journal of General Systems 17 (2-3) (1990) 191–209.
[46] I. Durbach, An empirical test of the evidential reasoning approach's synthesis axioms, Expert Systems with Applications 39 (12) (2012) 11048–11054.
[47] I. Durbach, T. Stewart, Using expected values to simplify decision making under uncertainty, Omega 37 (2) (2009) 312–330.
[48] I. Durbach, T. Stewart, An experimental study of the effect of uncertainty representation on decision making, European Journal of Operational Research 214 (2011) 380–392.
[49] I. Durbach, T. Stewart, A comparison of simplified value function approaches for treating uncertainty in multi-criteria decision analysis, Omega 40 (2012) 456–464.
[50] R. Dyckerhoff, Decomposition of multivariate utility functions in non-additive expected utility theory, Journal of Multi-Criteria Decision Analysis 3 (1) (1994).
[51] D. Ellsberg, Risk, ambiguity, and the savage axioms, Quarterly Journal of Economics 75 (1961) 643–669.
[52] Z. Fan, Y. Liu, B. Feng, A method for stochastic multiple criteria decision making based on pairwise comparisons of alternatives with random evaluations, European Journal of Operational Research 207 (2) (2010) 906–915.
[53] J. Figueira, S. Greco, M. Ehrgott (Eds.), Multiple Criteria Decision Analysis: State of the Art Surveys, Springer, 2005.
[54] B. Fischoff, W. de Bruin, U. Güvenc, D. Caruso, L. Brilliant, Analyzing disaster risks and plans: an avian flu example, Journal of Risk and Uncertainty 33 (2006) 131–149.
[55] P. Fishburn, Foundations of risk measurement. I. Risk as probable loss, Management Science 30 (4) (1984) 396–405.
[56] S. French, Uncertainty and imprecision: Modelling and analysis, Journal of the Operational Research Society 46 (1995) 70–79.
[57] P. Garthwaite, J. Kadane, A. O'Hagan, Statistical methods for eliciting probability distributions, Journal of the American Statistical Association 100 (470) (2005) 680–701.
[58] P. Garthwaite, A. O'Hagan, Quantifying expert opinion in the UK water industry: An experimental study, The Statistician (2000) 455–477.
[59] T. Gilovich, D. Griffin, D. Kahneman, Heuristics and biases: The psychology of intuitive judgment, Cambridge University Press, 2002.
[60] P. Goodwin, G. Wright, Decision Analysis for Management Judgement, forth ed., John Wiley & Sons, Chichester, 2009.
[61] M. Goovaerts, F. de Vylder, J. Haezendonck, Insurance Premiums, North-Holland, Amsterdam, 1984.
[62] S. Greco, M. Ehrgott, J. Figuera (Eds.), Trends in Multiple Criteria Decision Analysis, Springer, 2010.
[63] S. Greco, M. Inuiguchi, R. Slowinski, Fuzzy rough sets and multiple-premise gradual decision rules, International Journal of Approximate Reasoning 41 (2) (2006) 179–211.
[64] S. Greco, B. Matarazzo, R. Slowinski, A new rough set approach to multicriteria and multiattribute classification, in: Rough Sets and Current Trends in Computing, Springer, 1998, pp. 60–67.
[65] S. Greco, B. Matarazzo, R. Słowinski, Handling missing values in rough set analysis of multi-attribute and multi-criteria decision problems, in: New Directions in Rough Sets, Data Mining, and Granular-Soft Computing, Springer, 1999, pp. 146–157.
[66] S. Greco, B. Matarazzo, R. Slowinski, Rough approximation of a preference relation by dominance relations, European Journal of Operational Research 117 (1999) 63–83.
[67] S. Greco, B. Matarazzo, R. Słowiński, The use of rough sets and fuzzy sets in MCDM, in: T. Gal, T. Stewart, T. Hanne (Eds.), Multicriteria Decision Making: Advances in MCDM Models, Algorithms, Theory, and Applications, Kluwer Academic Publishers, Dordrecht, 1999 (chapter 12).
[68] S. Greco, B. Matarazzo, R. Slowinski, Rough sets theory for multicriteria decision analyis, European Journal of Operational Research 129 (2001) 1–47.
[69] S. Greco, B. Matarazzo, R. Słowiński, Dominance-based rough set approach to decision under uncertainty and time preference, Annals of Operations Research 176 (1) (2010) 41–75.
[70] J. Hadar, W. Russell, Rules for ordering uncertain prospects, American Economic Review 59 (1) (1969) 25–34.
[71] C. Huang, D. Kira, I. Vertinsky, Stochastic dominance rules for multiattribute utility functions, Review of Economic Studies 41 (1978) 611–615.
[72] C. Hwang, K. Yoon, Multiple Attribute Decision Making: Methods and Appllications, Springer Verlag, 1981.
[73] G. Islei, G. Lockett, P. Naudé, Judgemental modelling as an aid to scenario planning and analysis, Omega 27 (1) (1999) 61–73.
[74] E. Jacquet-Lagrèze, Modelling preferences among distributions using fuzzy relations, in: H. Jungermann, G. De Zeeuw (Eds.), Models and Experiments in Risk and Rationality, D. Reidel Publishers, Dordrecht, 1977, pp. 99–114.
[75] S. Jain, M. Khare, Construction of fuzzy membership functions for urban vehicular exhaust emissions modeling, Environmental Monitoring and Assessment 167 (2010) 691–699.
[76] R. Jensen, An alternative scaling method for priorities in hierarchical structures, Journal of Mathematical Psychology (Print) 28 (3) (1984) 317–332.
[77] J. Jia, J. Dyer, A standard measure of risk and risk-value models, Management Science 42 (12) (1996) 1691–1705.
[78] J. Johnson, A. Bruce, Calibration of subjective probability judgments in a naturalistic setting, Organizational Behavior and Human Decision Processes 85 (2) (2001) 265–290.
[79] D. Kahneman, P. Slovic, A. Tversky, Judgment Under Uncertainty: Heuristics and Biases, Cambridge University Press, 1982.
[80] A. Kaufmann, M. Gupta, Fuzzy Mathematical Models in Engineering and Management Science, Elsevier, 1988.
[81] D. Keefer, S. Bodily, Three-point approximations for continuous random variables, Management Science 29 (5) (1983) 595–609.
[82] R. Keeney, H. Raiffa, Decisions with Multiple Objectives: Preferences and Value Tradeoffs, Cambridge University Press, Cambridge, 1993.
[83] C. Kirkwood, Estimating the impact of uncertainty on deterministic multiattribute evaluation, Management Science 38 (6) (1992) 819–826.
[84] A. Korhonen, Strategic financial management in a multinational financial conglomerate: a multiple goal stochastic programming approach, European Journal of Operational Research 128 (2001) 418–434.
[85] P. Krokhmal, M. Zabarankin, S. Uryasev, Modeling and optimization of risk, Surveys in Operations Research and Management Science 16 (2) (2011) 49–66.
[86] M. Kwiesielewicz, The logarithmic least squares and the generalized pseudoinverse in estimating ratios, European Journal of Operational Research 93 (3) (1996) 611–619.
[87] J. Lee, H. Lee-Kwang, Comparison of fuzzy values on a continuous domain, Fuzzy Sets and Systems 118 (3) (2001) 419–428.
[88] R. Lempert, D. Groves, S. Popper, S. Bankes, A general analytic method for generating robust strategies and narrative scenarios, Management Science 52 (4) (2006) 514–528.
[89] L. Leung, D. Cao, On consistency and ranking of alternatives in fuzzy AHP, European Journal of Operational Research 124 (1) (2000) 102–113.
[90] R. Levary, K. Wan, An analytic hierarchy process based simulation model for entry mode decision regarding foreign direct investment, Omega 27 (6) (1999) 661–677.

[91] Y. Liu, Z. Fan, Y. Zhang, A method for stochastic multiple criteria decision making based on dominance degrees, Information Sciences 181 (19) (2011) 4139–4153.
[92] R. Luce, E. Weber, An axiomatic theory of conjoint, expected risk, Journal of Mathematical Psychology 30 (1986) 188–205.
[93] T. Marchant, The measurement of membership by comparisons, Fuzzy Sets and Systems 148 (2) (2004) 157–177.
[94] T. Marchant, The measurement of membership by subjective ratio estimation, Fuzzy Sets and Systems 148 (2) (2004) 179–199.
[95] T. Marchant, A measurement-theoretic axiomatization of trapezoidal membership functions, IEEE Transactions on Fuzzy Systems 15 (2) (2007) 238–242.
[96] J. Martel, G. d'Avignon, G. Couillard, A fuzzy outranking relation in multicriteria decision making, European Journal of Operational Research 25 (1986) 258–271.
[97] L. Mikhailov, A fuzzy programming method for deriving priorities in the analytic hierarchy process, The Journal of the Operational Research Society 51 (3) (2000) 341–349.
[98] L. Mikhailov, P. Tsvetinov, Evaluation of services using a fuzzy analytic hierarchy process, Applied Soft Computing Journal 5 (1) (2004) 23–33.
[99] J. Mingers, J. Rosenhead, Problem structuring methods in action, European Journal of Operational Research 152 (3) (2004) 530–554.
[100] J. Miyamoto, P. Wakker, Multiattribute utility theory without expected utility foundations, Operations Research (1996) 313–326.
[101] G. Montibeller, H. Gummer, D. Tumidei, Combining scenario planning and multi-criteria decision analysis in practice, Journal of Multi-criteria Decision Analysis 14 (2006) 5–20.
[102] A. Morton, B. Fasolo, Behavioural decision theory for multi-criteria decision analysis: a guided tour, Journal of the Operational Research Society 60 (2) (2009) 268–275.
[103] B. Munier, M. Machina (Eds.), Models and Experiments in Risk and Rationality, Kluwer Academic Publishers, Dordrecht, 1994.
[104] A. Murphy, R. Winkler, Credible interval temperature forecasting: some experimental results, Monthly Weather Review 102 (11) (1974) 784–794.
[105] I. Newby-Clark, M. Ross, R. Buehler, D. Koehler, D. Griffin, People focus on optimistic scenarios and disregard pessimistic scenarios while predicting task completion times, Journal of Experimental Psychology: Applied 6 (3) (2000) 171–181.
[106] M. Nowak, Preference and veto thresholds in multicriteria analysis based on stochastic dominance, European Journal of Operational Research 158 (2) (2004) 339–350.
[107] A. O'Hagan, J. Oakley, Probability is perfect, but we can't elicit it perfectly, Reliability Engineering and System Safety 85 (1-3) (2004) 239–248.
[108] J. Quiggin, A theory of anticipated utility, Journal of Economic Behavior and Organization 3 (4) (1982) 323–343.
[109] C. Ram, G. Montibeller, A. Morton, Extending the use of scenario planning and MCDA for the evaluation of strategic options, Journal of the Operational Research Society 62 (5) (2010) 817–829.
[110] T. Saaty, The Analytic Hierarchy Process, McGraw-Hill, New York, 1980.
[111] R. Sarin, M. Weber, Risk-value models, European Journal of Operational Research 70 (1993) 135–149.
[112] P. Schoemaker, Scenario planning: a tool for strategic thinking, Sloan Management Review 36 (1995). p. 25.
[113] G. Shafer, A Mathematical Theory of Evidence, Princeton University Press, 1976.
[114] J. Shanteau, Competence in experts: the role of task characteristics, Organizational Behavior and Human Decision Processes 53 (1992) 252–266.
[115] Y. Siskos, E. Grigoroudis, N. Matsatsinis, UTA methods, in: Figueira et al. [53], chapter 8.
[116] R. Slowinski, D. Vanderpooten, A generalized definition of rough approximations based on similarity, IEEE Transactions on Knowledge and Data Engineering 12 (2) (2000) 331–336.
[117] C. Spetzler, C.-A. Staël Von Holstein, Probability encoding in decision analysis, Management Science (1975) 340–358.
[118] C. Starmer, Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk, Journal of Economic Literature 38 (2000) 332–382.
[119] T. Stewart, Simplified approaches for multicriteria decision making under uncertainty, Journal of Multi-criteria Decision Analysis 4 (1995) 246–258.
[120] T. Stewart, Measurements of risk in fisheries management, ORiON 14 (1/2) (1998) 1–15.
[121] T. Stewart, Dealing with uncertainties in MCDA, in: Figueira et al. [53], chapter 11.
[122] T. Stewart, S. French, J. Rios. Integration of multicriteria decision analysis and scenario planning, in preparation.
[123] T. Stewart, F. Losa, Towards reconciling outranking and value measurement practice, European Journal of Operational Research 145 (3) (2003) 645–659.
[124] T. Tervonen, J. Figueira, A survey on stochastic multicriteria acceptability analysis methods, Journal of Multi-Criteria Decision Analysis 15 (1–2) (2008) 1–14.
[125] U. Thole, H. Zimmermann, P. Zysno, On the suitability of minimum and product operators for the intersection of fuzzy sets, Fuzzy Sets and Systems 2 (2) (1979) 167–180.
[126] E. Triantaphyllou, C. Lin, Development and evaluation of five fuzzy multiattribute decision-making methods, International Journal of Approximate Reasoning 14 (4) (1996) 281–310.
[127] S. Tsaur, T. Chang, C. Yen, The evaluation of airline service quality by fuzzy MCDM, Tourism Management 23 (2) (2002) 107–115.
[128] A. Tsoukias, From decision theory to decision aiding methodology, European Journal of Operational Research 187 (1) (2008) 138–161.
[129] A. Tversky, Discussion, in: D. Bell, H. Raiffa, A. Tversky (Eds.), Decision Making: Descriptive, Normative, and Prescriptive Interactions, Cambridge University Press, Cambridge, 1988, pp. 599–612.
[130] A. Tversky, D. Kahneman, Judgment under uncertainty: heuristics and biases, Science 185 (1974) 1124–1131.
[131] A. Tversky, D. Kahneman, Advances in prospect theory: cumulative representation of uncertainty, Journal of Risk and Uncertainty 5 (4) (1992) 297–323.
[132] K. Van der Heijden, Scenarios: The Art of Strategic Conversation, John Wiley & Sons, New York, 1996.
[133] J. van der Pas, W. Walker, V. Marchau, G. Van Wee, D. Agusdinata, Exploratory MCDA for handling deep uncertainties: the case of intelligent speed adaptation implementation, Journal of Multi-Criteria Decision Analysis 17 (1/2) (2010) 1–23.
[134] L. Vargas, Reciprocal matrices with random coefficients, Mathematical Modelling 3 (1) (1982) 69–81.
[135] J. Verkuilen, Assigning membership in a fuzzy set analysis, Sociological Methods & Research 33 (4) (2005) 462–496.
[136] J. von Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton University Press, New York, 1953.
[137] D. Von Winterfeldt, W. Edwards, Decision Analysis and Behavioural Research, Cambridge University Press, London, 1986.
[138] P. Wakker, D. Deneffe, Eliciting von Neumann–Morgenstern utilities when probabilities are distorted or unknown, Management Science (1996) 1131–1150.
[139] X. Wang, E. Kerre, Reasonable properties for the ordering of fuzzy quantities (I), Fuzzy Sets and Systems 118 (3) (2001) 375–385.
[140] X. Wang, E. Kerre, Reasonable properties for the ordering of fuzzy quantities (II), Fuzzy Sets and Systems 118 (3) (2001) 387–405.
[141] Y. Wang, K. Chin, An eigenvector method for generating normalized interval and fuzzy weights, Applied Mathematics and Computation 181 (2) (2006) 1257–1275.
[142] Y. Wang, T. Elhag, Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment, Expert Systems With Applications 31 (2) (2006) 309–319.
[143] Y. Wang, T. Elhag, Z. Hua, A modified fuzzy logarithmic least squares method for fuzzy analytic hierarchy process, Fuzzy Sets and Systems 157 (23) (2006) 3055–3071.
[144] Y. Wang, Y. Luo, Z. Hua, On the extent analysis method for fuzzy AHP and its applications, European Journal of Operational Research 186 (2) (2008) 735–747.
[145] Y. Wang, C. Parkan, Multiple attribute decision making based on fuzzy preference information on alternatives: ranking and weighting, Fuzzy sets and systems 153 (3) (2005) 331–346.
[146] Y. Wang, J. Yang, D. Xu, K. Chin, The evidential reasoning approach for multiple attribute decision analysis using interval belief degrees, European Journal of Operational Research 175 (1) (2006) 35–66.
[147] E. Weber, W. Bottom, An empirical evaluation of the transitivity, monotonicity, accounting and conjoint axioms for perceived risk, Organizational Behaviour and Human Decision Processes 45 (1990) 253–275.
[148] G. Whitmore, Third-degree stochastic dominance, American Economic Review 60 (3) (1970) 457–459.
[149] E. Wollenberg, D. Edmunds, L. Buck, Using scenarios to make decisions about the future: anticipatory learning for the adaptive co-management of community forests, Landscape and Urban Planning 47 (1–2) (2000) 65–77.
[150] G. Wright, P. Goodwin, Future-focussed thinking: combining scenario planning with decision analysis, Journal of Multi-criteria Decision Analysis 8 (1999) 311–321.
[151] D. Xu, An introduction and survey of the evidential reasoning approach for multiple criteria decision analysis, Annals of Operations Research 195 (1) (2012) 163–187.
[152] D. Xu, J. Yang, Y. Wang, The evidential reasoning approach for multi-attribute decision analysis under interval uncertainty, European Journal of Operational Research 174 (3) (2006) 1914–1943.
[153] R. Xu, Fuzzy least-squares priority method in the analytic hierarchy process, Fuzzy sets and systems 112 (3) (2000) 395–404.
[154] J. Yang, Rule and utility based evidential reasoning approach for multiattribute decision analysis under uncertainties, European Journal of Operational Research 131 (1) (2001) 31–61.
[155] J. Yang, Y. Wang, D. Xu, K. Chin, The evidential reasoning approach for mada under both probabilistic and fuzzy uncertainties, European Journal of Operational Research 171 (1) (2006) 309–343.
[156] J. Yang, D. Xu, On the evidential reasoning algorithm for multiple attribute decision analysis under uncertainty, IEEE Transactions on Systems, Man and Cybernetics, Part A 32 (3) (2002) 289–304.
[157] C. Yu, A GP-AHP method for solving group decision-making fuzzy AHP problems, Computers and Operations Research 29 (14) (2002) 1969–2001.
[158] H. Zank, Cumulative prospect theory for parametric and multiattribute utilities, Mathematics of Operations Research (2001) 67–81.
[159] K. Zaras, Rough approximation of a preference relation by a multi-attribute stochastic dominance for determinist and stochastic evaluation problems, European Journal of Operational Research 130 (2001) 305–314.

[160] K. Zaras, Rough approximation of a preference relation by a multi-attribute dominance for deterministic, stochastic and fuzzy decision problems, European Journal of Operational Research 159 (1) (2004) 196–206.
[161] K. Zaras, J. Martel, Multiattribute analysis based on stochastic dominance, in: Munier and Machina [103], pp. 225–248.
[162] Y. Zhang, Z. Fan, Y. Liu, A method based on stochastic dominance degrees for stochastic multiple criteria decision making, Computers & Industrial Engineering 58 (4) (2010) 544–552.