A simulation-based test of stochastic multicriteria acceptability analysis using achievement functions

Ian Durbach
Department of Statistical Sciences, University of Cape Town, South Africa

Keywords: multiple criteria analysis, goal programming

Abstract

Stochastic multicriteria acceptability analysis using achievement functions (SMAA-A) is a preference model for discrete-choice decision making that inverts the traditional goal programming process by asking what combinations of aspirations are necessary to make each alternative the preferred one, rather than what alternative is preferred given a set of aspirations. In this paper, we test the ability of the model to discern good-performing alternatives from poorly-performing ones using a simulation study. Simulation results show that a suitably detailed construction of the acceptability index is particularly important, and that the resulting model can be fruitfully applied in the selection of a shortlist of alternatives from a larger set with only very limited decision maker involvement.

1 Introduction

In 1976 economist Herbert Simon proposed the concept of 'satisficing' [6] as a cognitive model of human decision making. In multicriteria decision analysis terms, it was argued that decision makers search for an alternative that performs sufficiently well on all criteria without necessarily trying to maximise this performance. Although this process was not in itself prescriptive, it provided a cognitive basis for the operational research field of goal programming, which had already been in existence for some decades. The goal programming philosophy evaluates an alternative via the comparison of the performance of that alternative to an aspiration level or goal. Such a comparison naturally takes the form of a distance measure, so that an attempt is made to minimize the underachievement (as represented by a distance) between the aspiration and the performance of an alternative on each criterion.

The use of goal programming has subsequently been well-established as a normative model of human decision making. Nevertheless, the specification of behavioural interpretations for the set of aspirations remains an open research problem. The aspirations themselves are pieces of preference information that may be specified along a spectrum ranging from highly optimistic, even unachievable, values, to strict lower bounds on acceptable performance. Placing a set of aspirations in this spectrum may be difficult for even an experienced decision maker, and the cognitive processes guiding aspiration selection are poorly understood. This, in combination with the sensitivity of results to the choice of goals and the difficulty of selecting well-balanced goals, has resulted in most goal programming methods being applied explicitly or implicitly in an iterative manner. A number of so-called interactive procedures are available to guide the iterative reassessment of aspirations in order to comprehensively search the decision space (for example, see [9]). However, a similar problem exists in that little is known about the cognitive processes that determine how decision makers revise their aspirations. There is some evidence that they may terminate the revision process before exploring fully the options available to them [7].

Out of these problems, a new decision aid has arisen that inverts the interactive process to ask 'what combination of aspirations are required in order to make alternative i the preferred one?', rather than the usual line of questioning: 'what alternative is preferred given a set of aspirations?'. This methodology has been independently developed as *stochastic multicriteria acceptability analysis using achievement functions* (SMAA-A) in [4] and *acceptability analysis goal programming* (AAGP) in [1]. Although there are some practical differences, the methods are computationally very similar, and since becoming aware of the other work I have reverted to the former label. The method is based upon work done in the context of exploring the criterion weight space in value function and outranking methods [2, 5, 3], the so-called stochastic multicriteria acceptability analysis (SMAA) methods. The SMAA-A framework presents to the decision maker the consequences of certain preferences without demanding that any intra-criterion preference information be explicitly stated; its purpose is more exploratory than conventional decision aid.

In this paper, numerical results are provided in support of the SMAA-A model by testing via a simulation experiment the general ability of the model to provide an adequate shortlist of alternatives from a larger set. The paper is structured as follows. In the next section the decision problem is briefly described and key terms and notation are introduced. In section 3, the three-phase procedure of the SMAA-A method is presented. Section 4 describes the structure of the implemented simulation and discusses the results obtained. Finally, section 5 provides a conclusion to the paper.

2 Description of the decision problem

We refer to a set of n alternatives indexed by $i \in \{1, \ldots, n\}$ from which a choice must be made by the decision maker, and a set of m criteria indexed by $j \in \{1, \ldots, m\}$ to be used as a basis for evaluating the alternatives. Each criterion is associated with an underlying attribute, so that by z_{ij} we denote the evaluation of alternative i on criterion j. In what follows, criteria have been defined so that more is always preferred i.e. in a maximising sense. The aspiration level relating to the desired level of performance on criterion j is given by g_j, which can be assembled into a vector of aspirations $\mathbf{g} = [g_1, \ldots, g_m]$. The set of all possible aspiration vectors is given by $\mathbf{G}$, which then represents every possible combination of goals that can be specified by the decision maker. Given the set of attribute evaluations z_{ij} and aspirations g_j, it is possible to compute the deviation δ_{ij} of each evaluation from the corresponding goal on that criterion. Perhaps in contrast to conventional goal programming, we allow the deviational variables to be negative, in which case $-\delta_j$ becomes a measure of overachievement. Alternatives are rank ordered according to the minimization of a Wierzbicki scalarising function of the form

$$\Delta_i = \max_{j=1}^{m} w_j \delta_{ij} + \epsilon \sum_{j=1}^{m} w_j \delta_{ij} \tag{1}$$

where w_j is a weight indicating the relative importance (to be interpreted in a swing-weight sense) of criterion j, and ϵ is a small constant, say 0.02. For the purposes of the simulation, it is necessary to associate with every attribute evaluation z_{ij} a *value* to the decision maker $V(z_{ij})$, which we abbreviate as v_{ij}.

3 Stochastic multicriteria acceptability analysis using achievement functions

The SMAA-A model is structured as a three-phase process

1. Setting the upper and lower bounds on potential aspiration levels.
2. Computing a set of measures of acceptability for each alternative.
3. Synthesising the information contained in the acceptability measures in accordance with the aims of the decision aiding process.

3.1 Phase 1: Setting bounds on the aspiration levels

The starting point of the SMAA-A is a set of attribute evaluations z_{ij} and a set of criterion weights w_j. For conventional goal programming methods, it is then normal to set a single aspiration level on each attribute, that aspiration level indicating the desired level of performance. In contrast, for SMAA-A the decision maker is required to set two *aspiration bounds* on each attribute: a minimum i.e. maximally pessimistic, aspiration level, and a maximum i.e. maximally optimistic, aspiration level, so that the level of performance towards which the decision maker is striving should plausibly lie somewhere between these two bounds. The use of aspiration bounds releases the decision maker from the burden of specifying a single desired level of performance. For example, a particular decision maker might feel less comfortable making a statement such as 'I desire 50 units of attribute 1' than 'I would be very unhappy with any less than 20 units of that attribute, but I cannot reasonably expect more than 80 units'. Naturally the width between the bounds might narrow or widen as the decision maker becomes more or less aware of his or her preference structure.

We suggest that it is good practice to at least begin with little or no reference to the preferences of the decision maker, which can be achieved by setting the minimum and maximum aspiration levels to the minimum and maximum attribute evaluations on each criterion respectively. The aspirations resulting from such a state of complete uncertainty are termed *unbounded* aspirations. This approach effectively removes any behavioural connotations (and hence exposure to cognitive biases) from the aspiration levels, since they are not related to the preferences of the decision maker beyond the initial specification of the set of alternatives from which a selection is to be made. The upper and lower bounds on the aspiration levels may certainly be interactively narrowed at a later stage.

3.2 Phase 2: Computing measures of acceptability

As we have mentioned, the SMAA-A model is based on determining, for each alternative i, the set of aspirations $\mathbf{G}_i$ that makes alternative i better than any other alternative i.e. finding all $\mathbf{g}$ such that $\Delta_i \leq \Delta_k,\ \forall k,\ k \neq i$. When aspirations are unbounded, the non-existence of any such favourable aspiration vector $\mathbf{g}_i$ implies that alternative i is dominated and may be excluded from consideration. Since aspirations may be specified to an arbitrary precision, all other alternatives possess an infinite number of aspiration vectors that would make that alternative the preferred one. It then becomes necessary to use surrogate measures to describe the set of aspiration vectors associated with each alternative.

Acceptability Index

The collective set of aspiration vectors $\mathbf{G}_i$ forms a volume which may be calculated according to the (m-1)-dimensional integral

$$\text{vol}(\mathbf{G}_i) = \int_{\mathbf{G}_i} dg \tag{2}$$

The size of this volume indicates the variety of different aspirations that make alternative i the preferred alternative. Specifically, the volume can be scaled by the volume of all possible aspiration vectors to form the acceptability index

$$A(i) = \frac{\text{vol}(\mathbf{G}_i)}{\text{vol}(\mathbf{G})} \tag{3}$$

The acceptability index is the most important and easily interpreted measure of the general superiority of each alternative. Given the information provided by the decision maker when placing bounds on the aspiration levels, it represents the probability that alternative i is the best-ranked alternative. An alternative with a higher acceptability index is represented by a larger combination of aspirations. However, the acceptability index, based as it is only on the *best* i.e. first ranked, alternative, has the potential to be misleading. Most importantly, two or more alternatives located close to one another in the decision space decrease each other's acceptability indices, which may lead to a cluster of individually attractive alternatives being ignored in favour of a less attractive but more isolated alternative [3]. Such a problem, however, is easily avoided by a simple extension of the acceptability index taking into account information about other ranks. For example, we may compute the set of aspirations $\mathbf{G}_i^k$ that makes alternative i the k-ranked alternative, together with an associated volume $\text{vol}(\mathbf{G}_i^k)$ and acceptability index $A^k(i)$. In such a manner, we can combine the information about different ranks to obtain a more aggregated view of each alternative.

Ranges for favourable aspirations

For each alternative, the ranges of the favourable aspiration levels associated with each criterion can also be of interest. That is, each alternative has a (possibly empty) set of favourable aspiration vectors. By inspecting the set of vectors, we can compute the minimum aspiration on criterion j for which alternative i is still the preferred alternative (the minimum favourable aspiration), and the maximum aspiration on criterion j for which alternative i is the preferred alternative (the maximum favourable aspiration). These are relatively easily obtained, and show for each alternative the types of judgements that might exclude that alternative from being chosen. For example, upon reflection, a decision maker may feel that his or her aspiration levels for a particular criterion lay outside the range of the favourable aspiration levels.

Typical aspiration vectors

Bearing in mind that the sets of favourable aspiration vectors may be large, it may also become necessary to provide the decision maker with a single vector of aspirations that summarises in some sense the full set of favourable aspiration vectors. Several suggestions for implementing this idea are proposed in [2], including the use of the centre of gravity of the hypervolume, the vector with minimax distance to $\mathbf{G}_i$, and the centre of the ellipsoid inscribed in $\mathbf{G}_i$.

3.3 Phase 3: Synthesis of Information

The third phase of the model synthesises the available information and presents it to the decision maker in an appropriate way, signalling a return to the interaction between analyst and decision maker, relative to the fairly technical phases 1 and 2, which may proceed as described earlier with little or no decision maker input. In this context of providing useful decision support, SMAA-A can be employed in two broad guises: as a stand-alone methodology, or as a preprocessing tool for the generation of a shortlist from a larger set of alternatives.

SMAA-A as a stand-alone methodology

In using SMAA-A as a stand-alone MCDA methodology, attention is focused on the progression towards the choice of a preferred alternative, as well as on helping the decision maker to investigate the structure of the problem and his or her preference structure. The aim for SMAA-A is therefore to package the information in such a way that (a) it is easy for the decision maker to understand, and (b) is sufficient for the decision maker to arrive at a choice of alternative. In particular, the requirements of (b) imply that all three acceptability measures described in section 3.2 are likely to be required. The acceptability index $A(i)$ indicates the general breadth of aspiration levels that support alternative i without providing any information as to the nature of these aspiration levels. The ranges on favourable aspirations and the typical aspiration vector address this problem directly by presenting to the decision maker the types of judgements which would result in alternative i being the preferred alternative.

The ranges of favourable aspirations on any one criterion may overlap for some of the alternatives, so that it will not usually be feasible to *select* an alternative based on these results. However, they may be instrumental in *removing* an alternative from further consideration based on a particularly narrow or unsuitable range of favourable aspiration levels on one or more criteria. The typical aspiration vectors, on the other hand, represent far more concise information, and so may be more useful as a basis for the selection of a preferred alternative. However, it is important that both the decision maker and the analyst are aware that the vector is only ‘typical’, and that there may be more or less variability around this typical vector. This returns the synthesis process to the consideration of the acceptability index, which provides information on exactly this degree of variability associated with the selection of each alternative.

SMAA-A as a preprocessing tool

The SMAA-A methodology is suited to informing the decision maker about the *implications* of certain preferences, rather than about the nature of those preferences. As such there is nothing in the current method that will help the decision maker learn about his or her underlying preference structure or the available tradeoffs. By using SMAA-A as a preprocessing tool, the focus of the decision aiding shifts onto providing the decision maker with a shortlist of alternatives that are (a) attractive enough to be considered potentially optimal given the current state of preferential knowledge, and (b) are diverse in the sense that they possess different strengths and weaknesses. This frees the SMAA-A model from necessarily informing the decision maker about his or her own preferences, although naturally such information remains valuable. In particular, the acceptability index assumes primary importance, with the other measures of acceptability playing supporting roles by either providing additional information where the acceptability index is inconclusive, or by providing a cross-check by which to veto unsatisfactory alternatives.

Using the preference-free unbounded aspirations implies that all non-dominated alternatives possess a non-zero acceptability index, so that we have no prior idea as to which alternative may be suitable for selection. As the bounds on the aspirations are narrowed, the number of alternatives with non-zero acceptability indices decreases until, when aspirations on each criterion are represented by a single number, only one alternative is suitable. The main issue relating to the use of SMAA-A as a preprocessing tool is to what extent the upper and lower bounds on aspirations should be narrowed in order to ensure an acceptable shortlist of alternatives in terms of the two properties of attractiveness and diversity described earlier. It is this ability to generate a shortlist which we intend to investigate via simulation in the following section.

3.4 Comparison to the original SMAA-A methodology

The methodology presented above is computationally very similar to the original SMAA-A presented in [4], but differs conceptually in the application of these computations. The modifications to the original SMAA-A methodology and the consequences thereof are listed below:

1. Most importantly, the opportunity to include *some* preference information in the form of upper and lower bounds on aspiration levels is offered. Including preference information allows the SMAA-A method to situate itself along the continuum from preference-free to full-preference methods depending on the experience and confidence of the decision maker.

2. The inclusion of some preference information presents a new measure of acceptability: the ranges for favourable aspirations give the minimum and maximum aspiration on criterion j for which alternative i is still the preferred alternative.

3. We have stressed the use of the SMAA-A methodology in a 'preprocessing' capacity. That is, by starting from a preference-free position and incrementally adding small pieces of preference information, two benefits are achieved. Firstly, a shortlist containing the most favourable alternatives can be identified for closer inspection, perhaps by a different MCDA methodology. Secondly, the decision maker becomes slowly aware of his or her preference structure and the implications for the choice of an alternative.

4 A simulation-based investigation of the SMAA-A model

Experiments designed to respond to research questions in MCDA can generally take the form of either observations of real-world MCDA applications or simulation studies of conjectured problems. Where the research questions relate to the behavioural questions similar to those identified previously, a real-world experiment is in most circumstances the only approach capable of offering an answer. Other investigations into the long-run efficacy of a decision process may be better furthered by simulation studies. In this section we propose to test using simulation the general ability of the SMAA-A model to provide an adequate shortlist of alternatives from a larger set. Specifically, we attempt to answer the following questions on the basis that they are of fundamental importance to the practical usefulness of SMAA-A and are amenable to investigation via simulation:

- Is the SMAA-A model able to provide the decision maker with an attractive shortlist of alternatives?
- To what extent is preference information necessary or desirable i.e. does narrowing the upper and lower bounds on the aspiration levels influence the results?
- To what extent does the degree of interpolation between upper and lower bounds on the aspiration levels influence the results?

We aim to provide an initial indication of the potential of SMAA-A in general application, without necessarily exploring all aspects and eventualities of the method or problem context, leaving such investigation to future research.

4.1 The general simulation structure

All simulation studies in MCDA are faced with the difficult question of which aspects of the decision process to include as parameters of the model. Some of the detail of the real-world decision process should be excluded without oversimplifying the model to the point where the results are not clearly interpretable in the real-world. In short we want to abstract the model to a point where the results obtained are easily interpretable yet capable of making meaningful recommendations for real-world decision processes. The general simulation structure developed here comprises four elements

1. Creating a problem context comprising the attribute evaluations and criterion weights.
2. Generating an idealised 'true' rank ordering of alternatives using hypothesised value functions.
3. Applying the SMAA-A model and selecting a shortlist of alternatives based on the results of the model.
4. Evaluating the shortlist of alternatives with respect to the true rank ordering.

The problem context

We consider a decision problem consisting of n alternatives evaluated over m criteria. A shortlist of p alternatives is desired for closer consideration by the decision maker. The computation time required by the simulated SMAA-A model is exponential in the number of criteria considered and can be quite heavy for more than 8 criteria. This problem is the focus of current work making use of sampling techniques, but remains a problem for future research. Since the effect of the size of the decision problem is not the primary focus of the simulations, only four problem contexts are initially considered: $(n = 15, m = 5)$, $(n = 30, m = 5)$, $(n = 50, m = 5)$, and $(n = 50, m = 8)$. Thereafter the problem context is fixed at $(n = 50, m = 5)$, this being the most poorly-performing context out of the four considered. The evaluations z_{ij} are generated so as to lie on the unit hypersphere in order to guarantee that all alternatives are non-dominated. Without loss of generality we standardise each criterion to have a maximum attribute value of 1 and a minimum attribute value of 0. Criterion weights are uniformly generated in the interval $[0.2, 0.8]$ and standardised to sum to 1, ensuring that no attribute may be more than 4 times as important as another.

The idealised rank ordering of alternatives

The construction of an idealised rank order follows the structure used by Stewart [8], and is based on the assumption that the decision maker possesses an idealised underlying preference structure, which exists as a goal which is aimed towards even if its actual form is not consciously known. This idealised preference structure is assumed to satisfy the properties of completeness, transitivity, consistency, continuity, and preferential independence i.e. preferences may be represented by additive value functions. This conjecture should not be associated with the (behavioural) idea that underlying every DM is a additive value function. Rather it is an explication of the guiding nature of MCDA methods in the construction of a preference structure which accurately represents the decision maker. Most importantly, the idealised preference structure results in a 'true' rank ordering, which allows for the comparison of the rank ordering produced by this so-called true preference structure with rank orderings produced by the SMAA-A methodology.

We consider only concave value functions consisting of two linear segments, taking the form

$$v_j(z_{ij}) = \begin{cases} \frac{\lambda}{0.5} z_{ij} & \text{if } z_{ij} \leq 0.5 \\ \lambda + \frac{(1-\lambda)}{0.5}(z_{ij} - 0.5) & \text{if } z_{ij} > 0.5 \end{cases} \tag{4}$$

where λ is the value associated with the point of inflection at $z_{ij} = 0.5$, and is generated uniformly on the interval $[0.7, 0.9]$. The concave value function is closely related to goal programming and to the intuitive notion of aspiration setting in particular, where the values associated with attribute levels beyond the aspiration level are constant or increasing only slowly relative to below the aspiration level. This implies that a decision maker might well be satisfied with attribute levels that are even slightly above the inflection point, so that we might say with some degree of certainty that the aspirations of the decision makers are close to 0.5. In contrast, convex value functions imply only that a decision maker is likely to be unhappy unless attribute levels exceed 0.5 by *some* margin. For this reason we do not consider convex value functions in this simulation study. Given the form of the value functions, the true rank order is obtained by ordering the global values $V_i = \sum_{j=1}^{m} w_j v_{ij}$ from highest to lowest.

Applying the SMAA-A model

The simulation of the SMAA-A model requires that two parameters be specified: (1) upper and lower bounds on the aspiration levels, and (2) the size of the intervals used to interpolate between the upper and lower bounds. Table 1 shows the six combinations of upper and lower bounds that are used:

(((NOTE: Table 1 approximately here)))

We initially fix the lower and upper aspiration bounds at 0.5 and 0.9 respectively in order to investigate the effects of (a) different approximation levels by interpolating between the bounds using the intervals $0.0\dot{3}$, 0.05, 0.08, 0.10 and 0.20, and (b) changes to the numbers of alternatives n and criteria m considered. Thereafter, a constant interval of 0.10 and a problem context defined by 50 alternatives and 5 criteria is used for the other simulations. Once the parameters have been specified, the goal program resulting from each combination of aspiration levels is solved using (1) and the resulting rank order recorded. We consider only the acceptability index as a basis for rank ordering the alternatives; the other two measures are difficult to include in a simulation environment. Three different acceptability indices are considered: an index based only upon the first rank i.e. values of $a^1(i)$, an index based upon the first, second and third ranks i.e. on values of $\sum_{k=1}^{3} a^k(i)$, and an index based upon the first ten ranks i.e. values of $\sum_{k=1}^{10} a^k(i)$. The alternatives corresponding to the p highest values for the appropriate acceptability index are selected to form the shortlist of alternatives that is then presented to the decision maker for further evaluation.

Evaluating the results of the SMAA-A model

Given that the attribute values have been generated so that the top alternatives are generally extremely similarly valued, we limit our attention to the top three alternatives in the true rank order and their appearance in the shortlist. From the perspective of generating a shortlist that forms the input for a multicriteria *choice* problem, it is of little interest whether the true rank order is precisely replicated or even if, say, the fifthor seventh-best alternatives appear in the shortlist. The results of the simulations are presented in the form of the following five measures,

1. Pr(best): the number of simulations (out of 100) in which the true best alternative appeared in the shortlist.

2. Pr(2 of 3): the number of simulations (out of 100) in which two of the top three alternatives as measured by the true rank order appear in the shortlist.

3. Pr(3 of 3): the number of simulations (out of 100) in which all three top alternatives as measured by the true rank order appear in the shortlist.

4. Rank TB: the average position of the true best alternative in the rank order produced by the SMAA-A model.

5. Rank ShB: the average position in the true rank order of the best alternative as measured by the SMAA-A model.

4.2 Simulation results

Effect of intervals of interpolation

As might be expected, the amount of detail included in the interpolation between upper and lower aspiration bounds can exert a significant influence on the quality of results. The average positions of the true- and shortlisted-best alternatives in each other's rank orders are particularly adversely affected by insufficiently detailed interpolation; improvements of roughly 45% (2-2.5 ranks) are achieved when moving from the coarsest interpolation interval of 1/5 to the finest interval of 1/30. The other statistic that is badly affected is the probability of selecting the top three true-best alternatives in the shortlist of 10, which decreases from 84/100 to 51/100. The effect of the chosen interpolation interval is far less influential with regard to the other evaluation statistics; the probability that the true best alternative is included in the shortlist deteriorates only 15% from 94/100 to 81/100. Furthermore, much of the improvement in results is achieved without requiring that very small interpolation intervals be used. This is particularly meaningful if one considers that, for a problem comprising m criteria, halving the interpolation interval implies a 2^m-fold increase in the computing time required. Specifically, it appears as if, in the context of this simulation, an interpolation interval of 0.10 provides quite acceptable results on all fronts. The results of different interpolation intervals on the selection and ranking statistics are shown in figures 1 and 2.

(((NOTE: Figure 1 and Figure 2 approximately here)))

Problem context

Within the ranges considered the number of alternatives n and the number of critieria m exert only a moderate influence on results. The results are slightly complicated by the fact that the size of the shortlist p must clearly depend on the number of alternatives n. Therefore in table 2 we give the number of simulations (out of 100) in which the true best alternative appeared in the shortlists of $p = \{3, 5, 10\}$ alternatives, denoting this Pr(best)_p.

(((NOTE: Table 2 approximately here)))

Turning first to the effect of the number of alternatives included in the problem context, it is fairly encouraging to note that even when increasing the number of alternatives from 15 to 50, the true best alternative only slips 1.5 ranks in the rank order generated by the SMAA-A model. As a result the probability of locating the best alternative in a shortlist of 10 or 5 alternatives is left relatively unaffected by changes to the number of alternatives considered. The effect of the number of criteria considered by the decision problem is also relatively small, although the direction of the effect is perhaps surprising. Enlarging the problem context from 5 to 8 criteria results in an improvement of one rank in the position of the true best alternative in the SMAA-A rank order. A possible explanation for this behaviour is the greater differentiation of alternatives resulting from including more criteria, which decreases the extent to which neighbouring alternatives reduce each other's acceptability indices. As the magnitude of the effects of changes to the problem context is relatively small, and the general results remain promising for a range of problem contexts, we consider further only the problem context ($n = 50, m = 5$), it being the most poorly-performing of the four combinations considered.

Other effects

The choice of upper and lower bounds for aspiration levels can exert a fairly strong influence over the results, with a small narrowing of the bounds from the initial points $[0, 1]$ resulting in disproportionately large improvements in the quality of the shortlist. Table 3 shows the evaluations for different choices of upper and lower bounds.

(((NOTE: Table 3 approximately here)))

With the knowledge that the simulated decision maker's aspiration levels are in the vicinity of 0.5, it is evident that the performance of the SMAA-A model improves as the upper and lower bounds are narrowed around that vicinity. However, that result is of secondary importance; the main thrust of the simulation results is that the quality of the shortlist becomes excellent before a great deal of preference information is required to narrow the aspiration bounds to such an extent. If the bounds are narrowed only slightly to $[0.2, 0.9]$, the resulting shortlist contains two of the top three true best alternatives in 85% of the simulations. If one considers the information content associated with such a narrowing, it is apparent that it requires little more than an acknowledgement by the decision maker that *some* compromise is necessary. The movement of the lower bound from 0 to 0.2 requires almost no recourse to the preferences of the decision maker at all; well over half of the simulated alternatives would still fully satisfy such a combination of aspirations.

There is, however, a strong interaction between the bounds chosen and the form of the acceptability index, so that it is not particularly instructive to consider the effect of the bounds on their own. In particular, the narrowing of the bounds has a far lesser effect on the quality of the shortlist if a more comprehensive acceptability index is used. The interaction effect is shown in table 4, which gives the number of simulations in which the shortlist contains (a) the true best alternative, and (b) two of the top three true best alternatives.

(((NOTE: Table 4 approximately here)))

The implication of the results is that if a comprehensive acceptability index is constructed using the first ten ranks, then no real recourse to the decision maker's preferences is required. Using an acceptability index comprising ten terms, the true best alternative is included in the shortlist with greater than 90% probability without *any* involvement from the decision maker. However, if a simpler acceptability index is used, more emphasis is placed on the correct representation of preferences. This is evident from the relatively far greater improvements that are obtained by increasing the number of terms in the acceptability index when wide aspiration bounds have been set i.e. the cases of $[0, 1]$ and $[0.2, 1]$. In fact, when narrower bounds are used (indicating a greater reliance on the decision maker) the complexity of the acceptability index has little or no effect on the quality of the shortlist. There is therefore a tradeoff between the simplicity of the acceptability index and the dependence of the model on the input of the decision maker. Given that the number of terms making up the acceptability index can be increased without any further reference to the decision maker or substantially greater modelling effort, there seems to be no reason why the acceptability index should not be extended in this way. The previous result is reinforced in table 5 by the average ranking of the true best alternative in the shortlist, and the average ranking of the shortlist's best alternative in the true rank order.

In fact, it may be possible to work with a smaller shortlist than the ten-alternative list described in the previous results. The simulation results indicate that, where a comprehensive acceptability index comprising the first ten ranks is used, a shortlist of five alternatives contains the true best alternative in at least 73% of the simulations and two of the top three true best alternatives in at least 70%, while a shortlist of three alternatives contains the true best alternative in at least half of the simulations. Depending on the nature of the decision process, that is if the decision is relatively unimportant or repetitive so that bounds can be narrowed, these smaller shortlists might be used with some justification. Table 6 shows the number of simulations for which the best alternative appeared in the respective shortlists, when the most comprehensive acceptability index i.e. $k = 10$, is used.

(((NOTE: Table 5 and Table 6 approximately here)))

4.3 Implications for practice

The simulation results suggest that the SMAA-A methodology may be applied to obtain a consistently good shortlist of alternatives for closer consideration by the decision maker, provided that some care is taken in the application of the model. In particular, it is important that the acceptability index does not focus only on the first rank, but is extended to cover further positions in the rank order. In the simulation results the preferred acceptability index used as many ranks as made up the shortlist i.e. $p = 10$. However, some caution is required in extending the acceptability index to include more than a single term. Although a more well-rounded view of an alternative's performance is obtained by increasing the scope of the acceptability index, the focus should not drift towards mediocre or poor rankings. In doing so we may run the risk of losing the ability to distinguish the truly superior alternatives. An important point in interpreting the simulation results is that even the largest shortlist ($p = 10$) is small relative to the number of alternatives, so that an alternative must perform well to appear on the shortlist as a prospective 'preferred alternative'.

If the construction of the acceptability index is handled in this way, the effect of the choice of upper and lower bounds on aspiration levels is greatly reduced. This allows for the model to be applied with relatively wide aspiration bounds on the basis that this requires little contribution from the decision beyond a cursory examination. In most cases it is sufficient to merely fix the upper and lower aspiration bounds at the best and worst local attribute levels respectively, which can be done independently of the decision maker. Beyond these results, we have shown that the resolution provided by a simple numerical approximation to the multidimensional volumes is adequate in distinguishing the set of superior alternatives, so that the methodology may be implemented using nothing more sophisticated than a spreadsheet application. In fact, the effect on the results of the interpolation interval used is relatively less than the effects exercised by either the construction of the acceptability index or the choice of aspiration bounds.

5 Conclusions

The results obtained from the simulation experiment indicate that the SMAA-A method seems particularly well suited to narrowing a set of alternatives down to a smaller shortlist for closer consideration. In this role, the SMAA-A model may be used to eliminate a large proportion of (not obviously) unattractive alternatives with little or no direct involvement from the decision maker. In our simulations we were able to reduce a list of 50 alternatives to a shortlist of 10 while still retaining in the overwhelming majority of cases the best alternatives. The results suggest that SMAA-A could play a role in the development of a preprocessing stage of decision making, located between the traditional problem structuring and problem solving phases. This preprocessing phase has as its fundamental aims the simplification of the problem context to be considered, and the gradual exposure of the decision maker to his or her preferences and the tradeoffs on offer.

The most important aspect of applying the SMAA-A model appears to be ensuring that a sufficiently comprehensive acceptability index is used. Provided that such an acceptability index is used, the quality of the shortlisted alternatives remains excellent even if no further preference information is obtained from the decision maker i.e. the unbounded aspirations are used. However, narrowing the aspiration bounds in consultation with the decision maker, in conjunction with a suitably detailed acceptability index, may allow for a smaller shortlist to be used. Furthermore, the narrowing of the aspiration bounds would provide an important opportunity for the decision maker to begin to explore his or her preferences and the implications thereof.

Upper bound	1	1	1	0.9	0.9	0.75
Lower bound	0	0.2	0.5	0.2	0.5	0.45

Table 1: Upper and lower aspiration bounds used in the simulation study

Context	Rank TB	Rank ShB	$\Pr(\text{best})_{10}$	$\Pr(\text{best})_5$	$\Pr(\text{best})_3$
$(n = 15, m = 5)$	2.0	2.0	100	97	85
$(n = 30, m = 5)$	2.3	2.2	100	93	81
$(n = 50, m = 5)$	3.4	3.4	98	87	69
$(n = 50, m = 8)$	2.4	2.3	99	93	81

Table 2: Effect of number of alternatives n and criteria m on shortlist quality

goals	Rank TB	Rank ShB	Pr(best)	Pr(2 of 3)	Pr(3 of 3)
$[0, 1]$	12.2	9.3	56	52	27
$[0.2, 1]$	8.6	7.5	69	69	41
$[0.5, 1]$	3.9	4.7	95	94	68
$[0.2, 0.9]$	5.3	4.8	88	85	57
$[0.5, 0.9]$	3.4	3.4	97	98	85
$[0.45, 0.75]$	2.4	2.2	100	99	91

Table 3: Main effects of upper and lower aspiration bounds on shortlist quality

	Pr(best)			Pr(2 of 3)		
goals	$k = 1$	$k = 3$	$k = 10$	$k = 1$	$k = 3$	$k = 10$
$[0, 1]$	27	51	91	16	46	94
$[0.2, 1]$	47	69	90	44	69	95
$[0.5, 1]$	94	95	97	91	95	97
$[0.2, 0.9]$	76	89	98	69	88	99
$[0.5, 0.9]$	95	97	98	96	99	100
$[0.45, 0.75]$	99	100	100	99	99	100

Table 4: Effect of interaction between type of acceptability index and aspiration bounds

	Rank TB			Rank ShB		
goals	$k = 1$	$k = 3$	$k = 10$	$k = 1$	$k = 3$	$k = 10$
$[0, 1]$	20.7	11.8	4.1	14.1	9.8	4.1
$[0.2, 1]$	13.0	8.2	4.6	11.2	7.2	4.2
$[0.5, 1]$	4.5	3.8	3.4	5.8	4.9	3.3
$[0.2, 0.9]$	7.7	4.8	3.2	7.2	4.6	2.6
$[0.5, 0.9]$	3.5	3.2	3.5	3.9	3.1	3.1
$[0.45, 0.75]$	2.6	2.4	2.2	2.5	2.1	2.1

Table 5: Effects of interaction between aspiration bounds and the acceptability index on average rank statistics

goals	Shortlist of 10	Shortlist of 5	Shortlist of 3
$[0, 1]$	91	76	56
$[0.2, 1]$	90	73	53
$[0.5, 1]$	97	82	65
$[0.2, 0.9]$	98	83	67
$[0.5, 0.9]$	99	87	67
$[0.45, 0.75]$	100	96	85

Table 6: Number of simulations (out of 100) locating the true best alternative in different sized shortlists ($k = 10$)

References

[1] I. Durbach. Acceptability analysis in goal programming. *Department of Statistical Sciences Technical Report Series*, University of Cape Town, South Africa, 2003.

[2] R. Lahdelma, J. Hokkanen, and P. Salminen. SMAA – stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, 106:139–143, 1998.

[3] R. Lahdelma and P. Salminen. SMAA-2: stochastic multi-criteria acceptability analysis for group decision making. *Operations Research*, 49(3):444–454, 2001.

[4] R. Lahdelma, K. Miettinen and P. Salminen. Stochastic multicriteria acceptability analysis using achievement functions. *Turku Centre for Computer Science Technical Report Series*, 459, University of Turku, Finland, 2002.

[5] R. Lahdelma and P. Salminen. Pseudo-criteria versus linear utility functions in stochastic multi-criteria acceptability analysis. *European Journal of Operational Research*, 141:454–469, 2002.

[6] H. Simon. *Administrative Behaviour.* The Free Press, New York, 1976.

[7] T. Stewart. Experience with prototype multicriteria decision support systems for pelagic fish quota determination. *Naval Research Logistics*, 35:719–731, 1988.

[8] T. Stewart. Robustness of additive value function methods in MCDM. *Journal of Multi-criteria Decision Analysis*, 5:301–309, 1996.

[9] T. Stewart. Concepts of interactive programming. In T. Gal, T. Stewart, and T. Hanne, editors, *Multicriteria Decision Making: Advances in MCDM Models, Al-*

gorithms, Theory, and Applications, chapter 10. Kluwer Academic Publishers, Dordrecht, 1999.

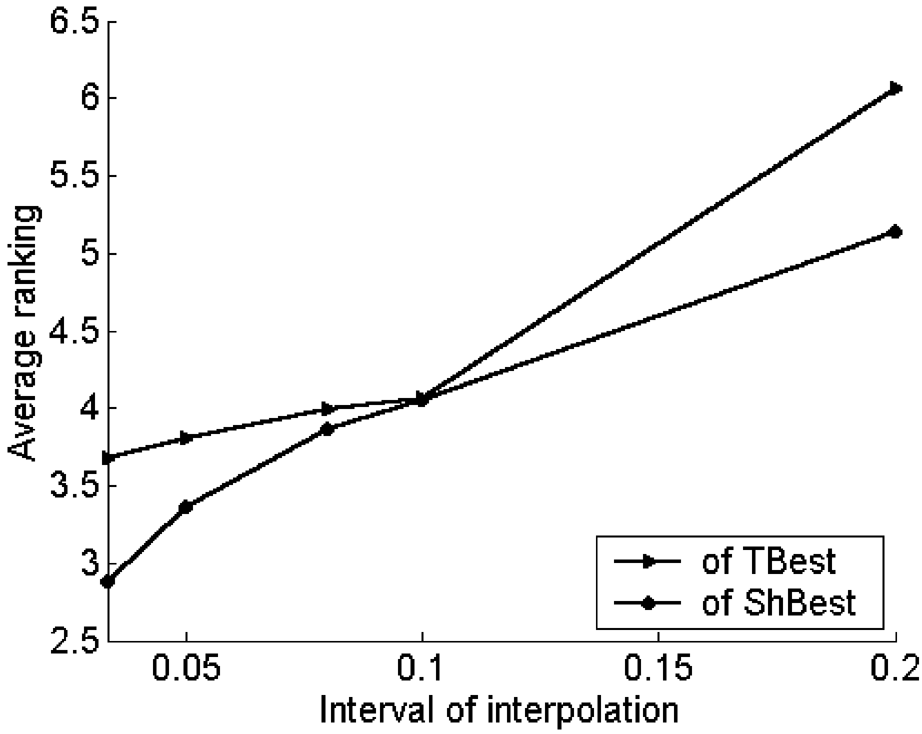

Figure 1: Effect of interpolation interval on probability of selection

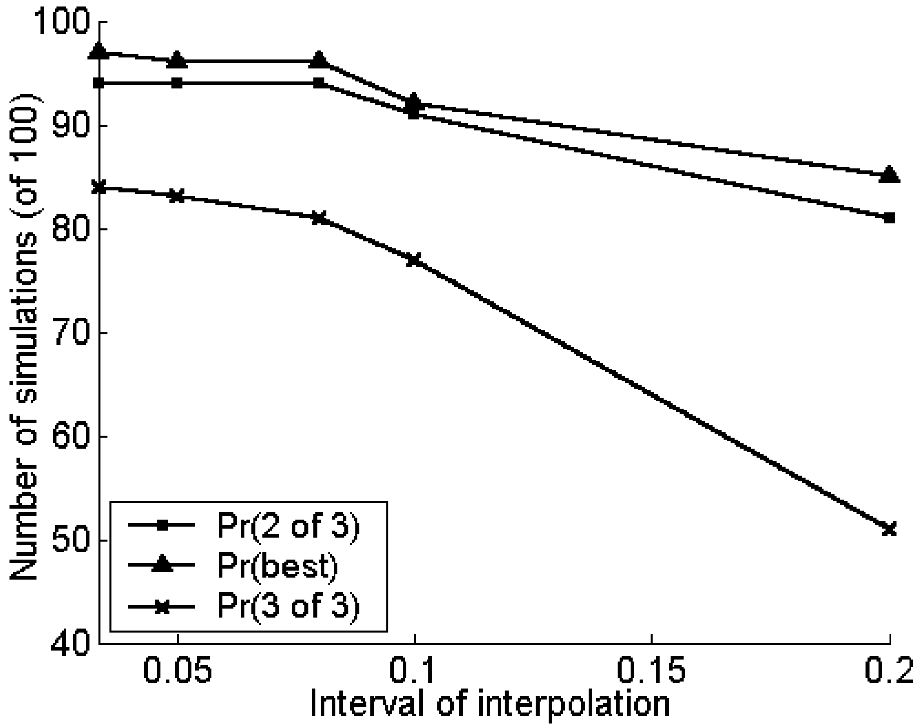

Figure 2: Effect of interpolation interval on average rankings