Multi-Criteria Decision Analysis on Aircraft Stringer Selection

Shuguang Song* You Ren† Paul d. Sampson‡

Abstract
Multi-Criteria Decision Analysis (MCDA) problems often involve multiple Decision Makers (DMs). In this paper, we present several decision analysis algorithms, considering both subjective and objective decision criteria with different strategies to account for uncertainty. We address the uncertainty and availability of weights for decision criteria, and develop probability scoring for the criteria. We demonstrate an application of our method with a case study concerning aircraft stringer decisions. This is sample text and needs to be completely replaced before submitting your paper. This is sample text and needs to be completely replaced before submitting your paper. This is sample text and needs to be completely replaced before submitting your paper. This is sample text and needs to be completely replaced before submitting your paper.

Key Words: multi-criteria decision analysis, Bayesian, scoring, uncertainty

1. Introduction

In aircraft manufacturing, a stringer is a thin strip of wood, metal, or carbon fiber to which the skin of the airplane is fastened. Consider the decision of choosing one of three stringer designs: Stringer I, Stringer II, and Stringer III. There are five decision criteria in this decision analysis problem. They are cost, cycle time per airplane, labor hours per airplane, rework rate and Technical Readiness Level (TRL). The cost of capital equipment for each of the three stringers is given. The criterion cycle time per airplane is used to measure the time needed to install the stringers per airplane. The criterion labor hours per airplane is used to measure total labor hours needed to install all stringers per airplane. Simulation models are developed for manufacturing processes of the three stringers according to their designs. The scale for TRL is from 1 to 10 and data were collected from 8 Decision Makers (DMs). The range of rework rate is from 0% to 100%. Ten DMs provided their subjective estimates of rework rate for the three stringers. Weights were also collected from two DMs for each of the five decision criteria. In this paper, we apply five Multi-Criteria Decision Analysis (MCDA) algorithms that involve multiple DMs by modeling the uncertainty in both types of decision factors and, in particular, uncertainty of weights, to rank the three stringer designs in aircraft manufacturing.

Multi-Criteria Decision Analysis (MCDA) involving multiple Decision Makers (DMs) has a broad applicability in finance, public policies, energy planning, nuclear waste, telecommunication network planning, and natural resources planning, see Figueira *et al.* (2005a), Hayashi (2000), Hajkowicz and Collins (2007), Brown (2009) and Brothers *et al.* (2009). Early applications were in military planning (Eckenrode, 1965). Figueira *et al.* (2005b) provided a thorough review of MCDA methods. Most MCDA methods study the deterministic consequences of alternatives on a set of criteria with sensitivity analysis. T. J. Stewart reviewed several methods addressing criteria value uncertainties in Figueira *et al.* (2005a). Decision criteria that follow distributions marginally and independently were studied in Hadar and Russell (1969). Keeney and Raiffa (1976) studied the case that utility function of every alternative has a distribution. Weight information is generally needed or elicited

*The Boeing Company, P.O. Box 3707, MS:4A-80, Tukwila, WA 98108, USA
†Department of Statistics, University of Washington, Seattle, WA 98195, USA
‡Department of Statistics, University of Washington, Seattle, WA 98195, USA

from DMs. Instead of asking for weights of criteria and reporting the single best alternative, Stochastic Multicriteria Acceptability Analysis (SMAA) explores the weight space by considering the uncertainty of weights, see Lahdelma *et al.* (1998), Lahdelma and Salminen (2001) and Tervonen and Figueira (2008). SMAA also takes into account the uncertainties of criteria values by imposing probability distributions. In the following, we highlight three aspects of our MCDA algorithms that involved in multiple DMs.

First, we categorize the decision criteria into two types: subjective decision criteria where values are subjective assessments from experts, and objective decision criteria where values come from historical numeric data or other sources such as simulation. To model the uncertainty of decision criteria, SMAA requires the specification of distributions. Lahdelma *et al.* (1998) assumed a uniform distribution for the criteria value. Lahdelma *et al.* (2006) and Lahdelma *et al.* (2009) assumed a multivariate normal distribution to handle the dependence among criteria and considered dependent uncertainties. In our proposed methodology, appropriate probability distributions or empirical distributions of objective criteria are estimated from data. To model the uncertainty of subjective criteria, we sample the data from a Bayesian posterior distribution. The Bayesian model can naturally handle the case of missing observations through prior, and it balances the prior and the observations obtained from DMs.

Second, we study the uncertainties that are embedded in the weights of criteria. The weights of criteria are usually elicited from the DMs and are very subjective. Similar to subjective criteria, we construct distributions for weights if weights are provided by multiple DMs. The Monte Carlo samples drawn from the weights distribution are able to fully capture the uncertainties in weights. If little or no weights information is available, we sample the weights from a uniform distribution, as in the SMAA approach.

Third, we show our Bayesian sampling frame work is compatible with other“winning” measures, such as pairwise winning index that was introduced by Leskinen *et al.* (2006). We call the winning measure a “probability score”. For a single criterion, “probability score” measures the probability that one alternative outperforms another.

The rest of this paper is organized as follows: uncertainty modeling is provided in Section 2. In Section 3, we present the scoring methods that are implemented in our decision analysis algorithms. In particular, we describe the probability scoring. A design case is used to illustrate the proposed decision analysis methods in Section 4. Finally, we provide some conclusions and discussion in Section 5.

2. Sampling Methods

Consider a MCDA problem that involves multiple DMs to evaluate m alternatives $\{A_1, A_2, \ldots, A_m\}$: there are n_1 subjective decision criteria $\{C_S^1, C_S^2, \ldots, C_S^{n_1}\}$ and n_2 objective decision criteria $\{C_O^1, C_O^2, \ldots, C_O^{n_2}\}$. For each subjective decision criteria C_S^j, where $j = 1, 2, \ldots, n_1$, its subjective assessment can be collected from a number of different DMs. Note that the number of DMs for each subjective decision criteria could be different. For each objective decision criteria C_O^j, the number of its data points could also be different. The total $n_1 + n_2$ weights for n_1 subjective decision criteria and n_2 objective decision criteria can be selected from a number of different experts. We also consider the case that there is no weight information in our methods. In the following, we provide several decision analysis algorithms to rank the m alternatives. The proposed methods for the multi-criteria decision making problems consist of two major parts: sampling and scoring. Sampling deals with uncertainty of criteria values and weights. For each sample, scoring method transforms the criteria values into utility score. The final total score, which is defined as the weighted sum of utility scores of each decision criteria, is computed for each sample.

Consequently, we obtain empirical distribution of total scores for all samples. For each sample, we can rank each of the total M alternatives from 1 to M. For all samples, we can count the number of an alternative being ranked rth choice for $r = 1, \ldots, M$. The rank acceptability index, which is defined as such counted number being divided by total number of samples, is also obtained for all M alternatives.

2.1 Sampling methods

In this section, we describe the sampling methods for subjective decision criteria, objective decision criteria, and weights.

2.1.1 Bayesian method

Subjective decision criteria can be categorical variables or continuous variables. For example, the decision criterion "Technical Readiness Level" is usually used to assess technological maturity of designs and its values are subjective input from DMs. The values of subjective decision criteria can also be continuous. Rework rate is usually a decision criteria to assess the percentage of parts that need to be reworked for each design and takes values from 0% to 100%. For a new design, the parts have not been massively manufactured and its rework rate is assessed by DMs. Thus, rework rate is a subjective decision criteria with continuous values.

Let X_i be the value for a subjective decision criterion from ith DM, where $i = 1, \ldots, n$. X_i takes l discrete values. Assume X_i follows a multinomial distribution:

$$X_i|\vec{p}\,\text{i.i.d.} \sim \text{Multinomial}(\vec{p}), i = 1, \ldots, d, \tag{1}$$

where $\vec{p} = (p_1, \ldots, p_l)$ and p_r is the probability of X_i taking rth discrete value. Further, we assume a prior distribution on the parameters $\vec{p}$:

$$\vec{p} \sim \text{Dirichlet}(\vec{\alpha} = 1, \ldots, 1), \tag{2}$$

where $\vec{\alpha}$ is a vector of length l. The Dirichlet distribution is the conjugate prior for a multinomial distribution. A symmetric Dirichlet distribution with common element values in the prior vector $\vec{\alpha}$ is often used as a non-informative prior, in which case no prior preference is placed on any support values. With values in $\vec{\alpha}$ set to 1, the symmetric Dirichlet distribution is equivalent to a uniform distribution over all points in its support. The data from experts can be reparameterized as $\vec{\beta} = (\beta_1, \ldots, \beta_l)$, where $\beta_r = \sum_{s=1}^{d} 1_{\{X_s=r\}}$. Bayesian analysis leads to the posterior distribution of $\vec{p}$ as follows,

$$\vec{p}|X_1, \ldots, X_d \sim \text{Dirichlet}(\vec{\gamma} = \vec{\alpha} + \vec{\beta}). \tag{3}$$

With the Bayesian model established, we propose the following 2-stage Bayesian sampling procedure:

For t from 1 to M,

Step 1. Sample one $\vec{p}_t$ from its posterior distribution $\vec{p} \sim \text{Dirichlet}(\vec{\gamma})$.
Step 2. Sample one data value X_t from its distribution $X_t \sim \text{Multinomial}(\vec{p}_t)$.

The samples for each C_S^j with discrete values can be obtained from this Bayesian 2-stage sampling method, where $j = 1, \ldots, n_1$.

For subjective decision criterion with continuous values such as rework rate, several DMs may provide their own estimates. But it is extremely challenging and unrobust to construct a continuous distribution based on only a handful of data points and then sample from the constructed continuous distribution. To utilize the handful of data points and overcome this challenge, we first discretize the continuous values and then apply the developed sampling procedure for categorical values. We categorize the range of the continuous variable into discrete "categories", $[a, b_1], (b_1, b_2], \ldots, (b_{l-1}, b_l]$, where l is the total number of categories, $a \leq b_1 \leq \ldots \leq b_l$ and $b_l = b$. Let X_i be the continuous variable. The discretization transforms the continuous value X_i into the categorical variable Y_i. The value X_i, $b_{r-1} < X_i \leq b_r$ is mapped to $Y_i = r$ and the support region of X_i, $[a, b]$, is mapped to the support region of Y_i, $\{1, 2, \ldots, l\}$. The same Bayesian model for a categorical variable described above can be applied to a discretized continuous variable. Both the variability in observations and in distribution parameters will be accounted for by the 2-stage sampling procedure: first sampling parameters from Bayesian posterior and then sampling criterion values based on the sampled parameters. Uniquely for continuous variables, a discretization and a conversion back to continuous values are added before and after the sampling procedure. The detailed sampling procedure for subjective continuous criteria is listed as follows:

For t from 1 to M,

Step 1. Sample one $\vec{p}_t$ from its posterior distribution $\vec{p} \sim$ Dirichlet$(\vec{\gamma})$.
Step 2. Sample one discrete value Y_t from its distribution $Y_t \sim$ Multinomial$(\vec{p}_t)$. Record the value $Y_t = r$.
Step 3. Sample one continuous value X_t from Uniform(b_{r-1,b_r}).

The samples for each C_S^j with continuous values can be obtained by this Bayesian 2-stage sampling method, where $j = 1, \ldots, n_1$.

2.1.2 Loss function method

We now consider a 1-stage sampling scheme using a single point estimate of the multinomial parameter vector $\vec{p}$. In Bayesian statistics, a loss function is defined as the expected loss with respect to the posterior distribution of parameters. The optimal parameter estimate is the one that minimizes the expected loss. For quadratic loss functions, the optimal estimate is the mean of the posterior distribution. In the case of posterior distribution of $\vec{p}$ that follows (3), we have $\hat{\vec{p}} = \frac{\vec{\gamma}}{1^T\vec{\gamma}}$. Thus, an alternative to the two-stage procedure, not accounting for uncertainty in estimation of the parameters of the multinomial expressed in the Dirichlet prior, is to use the posterior mean of the parameter vector $\vec{p}$. We then have a simpler 1-stage sampling procedure for criterion values as follows,

Obtain the posterior mean, $\hat{\vec{p}} = \frac{\vec{\gamma}}{1^T\vec{\gamma}}$. For t from 1 to M ,

Step 1. Sample one data value X_t from its distribution $X_t \sim$ Multinomial$(\hat{\vec{p}})$.

Compared to the 2-stage Bayesian sampling, the 1-stage sampling scheme for the loss function method has only one stage. In other words, the 1-stage sampling means to sample criterion values from a single multinomial distribution with its parameters $\vec{p}$ being the posterior mean of (3). As described before, the 1-stage sampling procedure can also be applied to continuous subjective decision criterion by discretization.

2.1.3 Sampling of weights

In this section, we first show how to study the uncertainty in the weights from a number of experts and then describe the method for the case that weights are not provided.

We first consider the case that weights for each decision criteria are collected from a number of DMs. Each DM assigns weights for n_1 subjective decision criteria and n_2 objective decision criteria. The values assigned on the weights can be constrained to integers from 1 to $n_1 + n_2$ with $n_1 + n_2$ being the most important criterion and 1 being the least important criterion. Note that ties and missing values are allowed. Since the weights will be normalized, i.e. $\sum_{j=1}^{n} w^j = 1$, the initial values of weight scale is not a concern. Let $(w_1^j, \ldots, w_d^j)$ denote the weight for the jth decision criterion by d DMs, where $j = 1, \ldots, n_1 + n_2$. The random variable W_i^j that is the weight by ith DM for C_S^j takes discrete values from $\{1, \ldots, n_1 + n_2\}$. Thus, the 2-stage sampling method and the 1-stage sampling method can both be applied the weight sampling.

It is known that the weights assigned to criteria play a critical role in the rankings of alternatives. In practice, DMs may be reluctant to provide the weights due to lack of knowledge about the relative consequences of different criteria. In the case that weights are not provided, Lahdelma *et al.* (1998) introduced Stochastic Multicriteria Acceptability Analysis (SMAA). The essential idea is to simulate weights uniformly from its space. The steps of sampling uniformly distributed normalized weights are described in Section 3.2 of Tervonen and Lahdelma (2007). We list the steps as follows to sample one vector of uniformly distributed normalized weights for c criteria.

Step 1. Simulate $c - 1$ random numbers from Uniform$(0, 1)$ distribution. Add 0 and 1 to form a vector of length $c + 1$.

Step 2. Sort the vector in ascending order, $(q_1, \ldots, q_{c+1})$.

Step 3. Compute the weights vector as, $w_i = q_{i+1} - q_i$, $i = 1, \ldots, c$.

It is proven by David (1970) that the procedure above would generate uniformly distributed normalized weights. Given the unconstrained weight space, every alternative is possible to be the best under a favorable set of criteria weights. For example, if an alternative has the highest score in criterion 1 and the lowest scores in all the other criteria, a weight vector of $(1, 0, \ldots, 0)$ makes this alternative rank first. Based on a large number of simulations of weights and criteria values, we can calculate the central (average) weight vector under which each alternative is ranked first respectively. The decision making is aided by choosing the most sensible set of weights that properly prioritizes the criteria. The alternative associated with the chosen weights is the best alternative. The central weight vector was also studied in the SMAA and SMAA-2, see Lahdelma *et al.* (1998) and Lahdelma and Salminen (2001).

3. Probability Scoring

Scoring is a process to map the decision criteria data to some numeric value through a so-called utility function, so that the score quantitatively measures which alternatives are better based on a set of criteria. We summarize the conventional methods of normalization tables and interval hull linear mapping, and then implement a probability scoring method accounting for uncertainties in criteria values and weights based on Bayesian 2-stage sampling.

A normalization table is usually provided by DMs after carefully studying the relationships among the criteria based on experience and thorough discussion. It explicitly lists the value range of each criterion and its associated score, see Taque (2005). The limitation of this method is that it does not consider the variation of input from different DMs. An interval hull linear mapping method was proposed by Tervonen *et al.* (2011). The key idea is to learn a mapping function from the criteria values observed, with some extreme observations removed. For alternative k on criterion j, the sampling procedure can yield an empirical sampling distribution for the jth criterion. A corresponding 95% confidence interval can then be obtained. The interval hull for criterion j is the smallest interval that contains 95% confidence intervals of all M alternatives on criterion j. We assume that there is a monotonic relationship between criterion values and scores. The two end points of the interval hull are mapped to the least and the most preferable values in the utility function, i.e., 0 and 1. The utility or mapping function is then assumed to be linear between the two end points. The interval hull based on 95% confidence intervals ensures the robustness of scores to outliers. The linear mapping is a simple way to implement and interpret. The DMs may apply a more sophisticated mapping function to the interval hull in order to tailor the method to a specific problem.

Here, we propose probability score to measure the probability that one alternative outperforms the other alternatives for a given decision criteria. Let X_1 denote the criterion variable for alternative 1, and X_2 denote the criterion variable for alternative 2. If the higher the criterion value, the better the alternative, the probability score of X_1 is then defined as $Pr(X_1 > X_2)$ and the score of X_2 is $Pr(X_2 > X_1)$. If the lower the criterion value, the better the alternative, the probability score of X_1 is then defined as $Pr(X_1 < X_2)$ and the score of X_2 is $Pr(X_2 < X_1)$. If the criterion is a categorical variable, the probability score is adjusted by adding half of the probability of X_1 equaling to X_2, i.e. $Pr(X_1 > X_2) + 0.5Pr(X_1 = X_2)$. In the follow, we describe the proposed probability scoring methods for both subjective decision criteria and objective decision criterion. Note that the algorithm below is based on the Bayesian sampling procedure as described before.

For one subjective criterion:

For sample t from 1 to M,

Step 1. Sample one parameter for Alternative 1, $\vec{p}_t^1$, from its posterior distribution; Sample one parameter for Alternative 2, $\vec{p}_t^2$, from its posterior distribution.
Step 2. Therefore, Criterion value of Alternative 1, X_t^1, follows Multinomial$(\vec{p}_t^1)$; Criterion value of Alternative 2, X_t^2, follows Multinomial$(\vec{p}_t^2)$; Calculate $Pr(X_t^1 < X_t^2)$. If the preference direction is increasing, probability score $s_{12} = Pr(X_t^1 > X_t^2)$. If the preference direction is decreasing, probability score $s_{12} = Pr(X_t^1 < X_t^2)$. Repeat the pairwise comparison for any two alternatives (total m alternatives).
Step 3. Summarize all pairwise comparison scores in a matrix

$$\begin{bmatrix} 0 & s_{12} & \cdots & s_{1m} \\ s_{21} & 0 & \cdots & s_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ s_{m1} & s_{m2} & \cdots & 0 \end{bmatrix}.$$

Calculate the row sums, then divided by $(m-1)$. $\bar{s}_k = \sum_{l=1}^{m} s_{kl}/(m-1)$, which represents the average probability of alternative k outperforming the others. Thus, we obtain a vector of probability scores $(\bar{s}_1, \ldots, \bar{s}_m)$ for each alternative on one subjective criterion under sample t.

For objective criteria, we bootstrap the observed values to get another sample, denoted as Sample t. We then calculate the pairwise probability score $Pr(X_t^1 > X_t^2)$ by convolution over empirical distributions of X_t^1 and X_t^2. Note if the sample size for an objective criterion is large, bootstrapping may not be necessary. The probability scores calculated from the original sample will be very close to the scores calculated from bootstrapped samples. We can save some computing time by computing the scores based on the original sample only.

The probability score derives from the probability of one alternative outperforming the other. Its value is from 0 to 1, with 1 meaning a winner, 0 meaning a loser and 0.5 meaning the two alternatives having equal performance. We can rescale the original probability scores from the support $[0,1]$ to a new scale $[-1,1]$. In the new score scale, 1 means a complete winner, -1 means a complete failure and 0 means that the two alternatives have the same performance.

4. Application of MCDA Algorithms to Aircraft Stringers

As described before, there are five decision criteria in the aircraft stringers selection problem. They are cost, cycle time per airplane, labor hours per airplane, rework rate and Technical Readiness Level (TRL). Rework rate and TRL are two subjective decision criteria. Simulation models are developed for manufacturing processes of the three stringers according to their designs. In this case study, we have 200 simulation runs for the cycle time per airplane, and 100 simulation runs for the labor hours per airplane. Note that missing values are allowed in our algorithms. Ties in weights are also allowed for decision criteria, see Table 1. A normalization table is provided in Table 2.

Decision Factor	weight from DM1	weight from DM2
cost	5	5
cycle time per airplane	1	2
labor hours per airplane	2	2
rework rate	3	4
TRL	4	3

Table 1: Weights of Decision Criteria.

Score	10	9	8	7	6	5	4	3	2	1
cost	0-50	50-80	80-90	90-100	100-110	110-120	120-130	130-140	140-150	150-160
TRL	10	9	8	7	6	5	4	3	2	1
rework rate	0%-10%	10%-20%	20%-30%	30%-40%	40%-50%	50%-60%	60%-70%	70%-80%	80%-90%	90%-100%
labor hours	0.5K-1.0K	1.0K-1.5K	1.5K-2.0K	2.0K-2.5K	2.5K-3.0K	3.0K-3.5K	3.5K-4.0K	4.0K-4.5K	4.5K-5.0K	5.0K-5.5K
cycle time	100-200-	200-300-	300-400-	400-500-	500-600-	600-700-	700-800-	800-900-	900-1K-	1K-1.1K

Table 2: Normalization Table for Stringer Study.

With weights, we provide five decision analysis algorithms. A combination of two sampling procedures (2-stage sampling and 1-stage sampling) and two scoring methods (normalization table and interval hull linear mapping method) yields four decision analysis algorithms. The fifth algorithm is based on probability scoring method that uses the Bayesian 2-stage sampling method for subjective decision criteria. For each decision analysis algorithm, 50,000 samples of criteria values and weights are generated from corresponding Bayesian 2-stage sampling procedure or 1-stage sampling procedure. Using the corresponding scoring method in each algorithm, we have 50,000 total weighted scores.

Algorithm	Stringer	Mean	95% CI
1-stage sampling, normalization table	Stringer I	7.03	(5.44, 8.21)
	Stringer II	7.56	(6.06, 8.54)
	Stringer III	7.99	(6.47, 8.93)
2-stage sampling, normalization table	Stringer I	7.02	(5.38, 8.21)
	Stringer II	7.56	(6.06, 8.56)
	Stringer III	7.99	(6.47, 8.93)
1-stage sampling, interval hull linear score	Stringer I	0.34	(0.12, 0.57)
	Stringer II	0.64	(0.43, 0.80)
	Stringer III	0.72	(0.51, 0.89)
2-stage sampling, interval hull linear score	Stringer I	0.34	(0.12, 0.56)
	Stringer II	0.64	(0.43, 0.80)
	Stringer III	0.72	(0.51, 0.89)
probability scoring	Stringer I	-0.47	(-0.73,-0.19)
	Stringer II	0.03	(-0.07, 0.13)
	Stringer III	0.44	(0.16, 0.69)

Table 3: Mean and 95% Confidence Interval of Total Weighted Scores.

Algorithm	Stringer	Rank 1	Rank2	Rank3
1-stage sampling, normalization table	Stringer I	0.07	0.24	0.69
	Stringer II	0.25	0.54	0.21
	Stringer III	0.68	0.22	0.10
2-stage sampling, normalization table	Stringer I	0.08	0.24	0.68
	Stringer II	0.24	0.54	0.22
	Stringer III	0.68	0.22	0.10
1-stage sampling, interval hull linear score	Stringer I	0.01	0.04	0.95
	Stringer II	0.25	0.72	0.03
	Stringer III	0.74	0.24	0.02
2-stage sampling, interval hull linear score	Stringer I	0.01	0.04	0.95
	Stringer II	0.26	0.71	0.03
	Stringer III	0.74	0.25	0.01
probability scoring	Stringer I	0.00	0.00	1.00
	Stringer II	0.01	0.99	0.00
	Stringer III	0.99	0.01	0.00

Table 4: Rank Accpetability Index for the Five Decision Analysis Algorithms.

Table 3 shows the means and 95% confidence intervals (CIs) of total weighted scores for the three stringers from each of the five decision algorithms. The higher the total weighted score, the better the design. All five algorithms show that Stringer III is the best design, Stringer II is the second best design, Stringer III is the third best design. Table 4 shows the rank acceptability index for the three stringers from each of the five algorithms. “Rank 1” column shows the probability of each stringer to be the best design. “Rank 2” column shows the probability of each stringer to be the second best design. “Rank 3” column shows the probability of each stringer to be the third best design. Again, all five algorithms yield a set of consistent results as shown in Table 3.

If weights are not available, we will use the sampling methods for weights in Section 2.1.3. The decision analysis is taken by choosing the most sensible set of central weights. Using the 1-stage sampling procedure and normalization Table 2, we choose the central weight vector with heavy weights on cost, cycle time per airplane, and labor hours per airplane. Our study shows that Stringer III is still the best design.

5. Conclusion and Discussion

Motivated by a real and typical decision problem in airplane manufacturing industry, we separate decision criteria by their nature into subjective and objective criteria, and treat them differently in the construction of the sampling distribution. All the uncertainties in criteria values, weights and uncertainties in sampling distribution parameters are considered in one coherent model through a sampling step and a scoring step. We explored in total five algorithms by combining the Bayesian 1-stage and 2-stage sampling procedures with normalization table and interval hull linear mapping for scoring, and embedding a pairwise winning index with the Bayesian 2-stage sampling scheme as our "probability scoring" method. Note that the "probability scoring" idea is related to the pairwise winning index introduced by Leskinen *et al.* (2006). Our contribution here is to implement this general pairwise index idea in this decision problem setting, where the nature of criteria, uncertainty of subjective criteria values from multiple DMs, uncertainty of weights, uncertainty of sampling distribution parameters are all considered under the unified Bayesian sampling and scoring approach. In our study of three aircraft stringers, the five algorithms and two metrics (total weighted score and rank acceptability index) give the consistent rankings for the three stringers. However, it is likely that the five algorithms can give different results for some other decision analysis problems. Thus, some extensive simulation study is needed to compare the performance of the given MCDA algorithms, and it may be worth the development of some ensemble MCDA algorithm.

ACKNOWLEDGMENT

This research was funded by The Boeing Company.

REFERENCES

Brothers, A. J., Mattigod, S. V., Strachan, D. M., Beeman, G. H., Kearns, P. K., Papa, A., and Monti, C. (2009). Resource-limited multiattribute value analysis of alternatives for immobilizing radioactive liquid process waste stored in Saluggia, Italy. *Decision Analysis*, **6**, 98–114.

Brown, R. V. (2009). Decision analysis. *Management Science*, **6**, 14–24.

David, H. A. (1970). *Order Statistics*. New York, John Wiley and Sons.

Eckenrode, R. T. (1965). Weighting multiple criteria. *Management Science*, **12**, 180–192.

Figueira, J., Mousseau, V., and Roy, B. (2005a). *ELECTRE methods. In: Figueira J, Greco S, Ehrgott M (eds) Multiple criteria decision analysis: state of the art surveys.* Berlin Heidelberg New York, Springer.

Figueira, J., Greco, S., and Ehrgott, M. (2005b). *Multiple criteria decision analysis: state of the art surveys*. New York, Springer.

Hadar, J. and Russell, W. R. (1969). Rules for ordering uncertain prospects. *The American Economic Review*, **59**, 25–34.

Hajkowicz, S. and Collins, K. (2007). A review of multiple criteria analysis for water resource planning and management. *Water Resources Management*, **21**, 1553–1566.

Hayashi, K. (2000). Multicriteria analysis for agricultural resource management: A critical survey and future perspectives. *European Journal of Operational Research*, **122**, 486–500.

Keeney, R. L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. New York, John Wiley and Sons.

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

Lahdelma, R., Hokkanen, J., and Salminen, P. (1998). SMAA - stochastic multiobjective acceptability analysis. *European Journal of Operational Research*, **106**, 137–143.

Lahdelma, R., Makkonen, S., and Salminen, P. (2006). Multivariate gaussian criteria in SMAA. *European Journal of Operational Research*, **170**, 957–970.

Lahdelma, R., Makkonen, S., and Salminen, P. (2009). Two ways to handle dependent uncertainties in multicriteria decision problems. *Omega*, **37**, 79–92.

Leskinen, P., Viitanen, J., Kangas, A., and Kangas, J. (2006). Alternatives to incorporate uncertainty and risk attitude in multicriteria evaluation of forest plans. *Forest Science*.

Taque, N. R. (2005). *The Quality Toolbox*. ASQ Quality Press, 2nd edition.

Tervonen, T. and Figueira, J. R. (2008). A survey on stochastic multicriteria acceptability analysis methods. *Journal of Multi-criteria Decision Analysis*, **15**, 1–14.

Tervonen, T. and Lahdelma, R. (2007). Implementing stochastic multicriteria acceptability analysis. *European Journal of Operational Research*, **178**, 500–513.

Tervonen, T., Valkenhoef, G., Buskens, E., Hillege, H. L., and Postmus, D. (2011). A stochastic multicriteria model for evidence-based decision making in drug benefit-risk analysis. *Statistics in Medicine*, **30**, 1419–1428.