On confident outrankings with multiple criteria of uncertain significance

Raymond Bisdorff
University of Luxembourg, FSTC/CSC/ILIAS
raymond.bisdorff@uni.lu

Abstract. When modelling preferences following the outranking approach, the sign of the majority margins do sharply distribute validation and invalidation of pairwise outranking situations. How can we be confident in the resulting outranking digraph, when we acknowledge the usual imprecise knowledge of criteria significance weights and a small majority margin? To answer this question, we propose to model the significance weights as random variables following more less widespread distributions around an average weight value that corresponds to the given deterministic weight. As the bipolarly valued random credibility of an outranking statement results from a simple sum of positive or negative independent and similarly distributed random variables, we may apply the CLT for computing likelihoods that a given majority margin is indeed positive, respectively negative.

Keywords: Multiple criteria decision aid; Uncertain criteria weights; Stochastic outranking relations; Confidence of the Condorcet outranking digraph.

Introduction

In a social choice problem concerning a very important issue like amending a country's Constitution, the absolute majority of voters is often not seen as sufficient for supporting a convincing social consensus. A higher majority of voters, two third or even three forth of them, may be required to support the bill in order to take effective decisions. Sometimes, even unanimity is required; a condition that, however, may generate in practice many indecisive situations. A similar idea is sometimes put forward in multiple criteria decision aiding in order to model global compromise preferences when the significance of the criterion are not known with sufficient precision. In his seminal work on the Electre I method (Roy [1], concerning a best unique choice problematique, Roy is clearly following this line of thought by proposing to choose a sufficiently qualified majority of criterial support before considering an outranking statement to be significant.

Following the SMAA approach (Tervonen et al. [2]), we are here proposing a different approach. The individual criteria significance weights are considered to be random variables. The bipolarly valued characteristic of the pairwise outranking situations (Bisdorff [3, 4]) appear hence to be sums of random variables of which we may assess the apparent likelihood of obtaining a positive weighted majority margin for each outranking situation. And depending on the seriousness of the decision issue, we may hence recommend to accept only those outranking statements that show a sufficiently high likelihood of 90% or 95%, for instance. We could also, in the limit accept only those statements which appear to be certainly supported by a weighted majority of criterial significance.

The paper is structured as follows. A first section is concerned with how to model the uncertainty we face for assessing precise numerical criteria significance weights. The second section illustrates how the likelihood of outranking situations may be estimated. The third section introduces the concept of confidence level of the valued outranking digraph, followed by short last section devoted to an illustrative example of confident best choice recommendation.

1 Modelling uncertain criteria significances

We have already extensively discussed some time ago (see Bisdorff [5]) the operational difficulty to numerically assess with sufficient precision the actual significance that underlies each criterion in a multiple criteria decision aid problem. Even, when considering that all criteria are equi-significant, it is not clear how precisely (how many decimals ?) such a numerical equality should be taken into account when computing the outranking characteristic values. In case of unequal significance of the criteria, it is possible to explore the stability of the Condorcet digraph with respect to the ordinal criteria significance structure (Bisdorff [6, 7]). One may also use indirect preferential observations for assessing via linear programming computations apparent significance ranges for each criterion (Dias [8]).

Here, we propose instead to consider the significance weights of a family F of n criteria to be independent random variables W_i, distributing the potential significance weights of each criterion $i = 1, ..., n$ around a mean value $E(W_i)$ with variance $V(W_i)$.

Choosing a specific stochastic model of uncertainty may be application specific. In the limited scope of this paper, we will illustrate the consequence of this design decision on the resulting outranking modelling with four slightly different models for taking into account the uncertainty with which we know the numerical significance weights: *uniform*, *triangular*, and two models of *Beta* laws, one more widespread and, the other, more concentrated. When considering that the potential range of a signficance weight is distributed between 0 and two times its mean value, we obtain the following random variates:

1. A continuous *uniform* distribution on the range 0 to $2 * E(W_i)$. Thus $W_i \sim \mathcal{U}(0, 2E(W_i))$ and $V(W_i) = \frac{1}{3}E(W_i)^2$;
2. A symmetric $beta(a,b)$ distribution with, for instance, parameters $a = 2$ and $b = 2$. Thus, $W_i \sim \mathcal{B}eta(2,2) \times 2E(W_i)$ and $V(W_i) = \frac{1}{5}E(W_i)^2$.
3. A symmetric *triangular* distribution on the same range with mode $E(W_i)$. Thus $W_i \sim \mathcal{T}r(0, 2E(W_i), E(W_i))$ with $V(W_i) = \frac{1}{6}E(W_i)^2$;
4. A narrower $beta(a,b)$ distribution with for instance parameters $a = 4$ and $b = 4$. Thus $W_i \sim \mathcal{B}eta(4,4) \times 2E(W_i)$, $V(W_i) = \frac{1}{9}E(W_i)^2$

It is worthwhile noticing that these four uncertainty models all admit the same expected value, $E(W_i)$, however, with a respective variance which goes decreasing from 1/3, to 1/9 of the square of $E(W_i)$ (see Fig. 1).

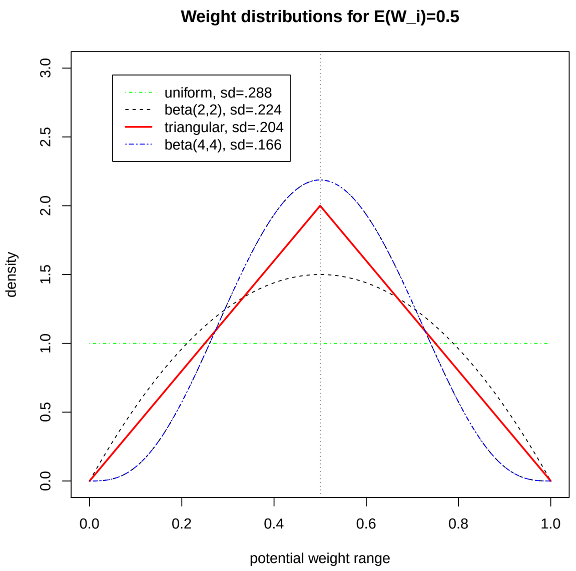

Figure 1. Four models of uncertain significance weights

We will limit in the sequel our attention to the triangular random model and explore now, without loss of generality, the resulting uncertainty we are going to model into the valued outranking digraph.

2 Likelihood of "*at least as good as*" situations

Let $A = \{x, y, z, ...\}$ be a finite set of n potential decision actions, evaluated on $F = \{1, ..., m\}$, a finite and coherent family of m performance criteria. On each criterion i in F, the decision actions are evaluated on a real performance scale $[0; M_i]$, supporting an upper-closed indifference threshold ind_i and a lower-closed preference threshold pr_i such that $0 \leqslant ind_i < pr_i \leqslant M_i$. The marginal performance of object x on criterion i is denoted x_i. Each criterion i is thus characterizing

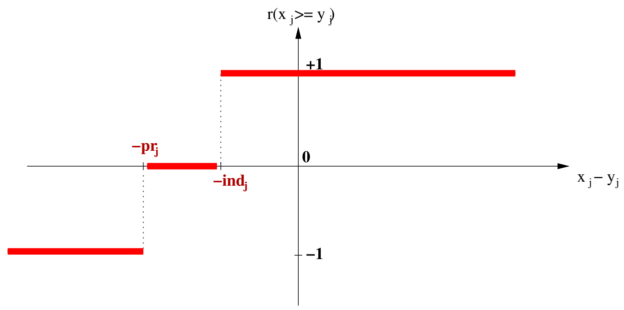

Figure 2. Characteristic function of marginal "*at least as good as*" statement

a marginal *double threshold* order $\succcurlyeq_i$ on A (see Fig. 2):

$$r(x \succcurlyeq_i y) = \begin{cases} +1 & \text{if} \quad x_i - y_i \geqslant -ind_i \\ -1 & \text{if} \quad x_i - y_i \leqslant -pr_i \\ 0 & \text{otherwise.} \end{cases} \qquad (1)$$

$+1$ signifies x is *performing at least as good as* y on criterion i,

-1 signifies that x is *not performing at least as good as* y on criterion i.

0 signifies that it is *unclear* whether, on criterion i, x is performing at least as good as y.

Each criterion $i \in F$ contributes the random significance W_i of his "*at least as good as*" characterization $r(\succcurlyeq_i)$ to the global characterization $\tilde{r}(\succcurlyeq)$ in the following way:

$$\tilde{r}(x \succcurlyeq y) \; = \; \sum_{i \in F} \left[\, W_i \cdot r(x \succcurlyeq_i y) \, \right] \qquad (2)$$

Thus, $\tilde{r}(x \succcurlyeq y)$ becomes a simple sum of positive or negative independent random variables with known means and variances where $\tilde{r} > 0$ signifies x is *globally performing at least as good as* y, $\tilde{r} < 0$ signifies that x is *not globally performing at least as good as* y, and $\tilde{r} = 0$ signifies that it is *unclear* whether x is globally performing at least as good as y.

From the Central Limit Theorem (CLT), we know that such a sum (Eq. 2) leads, with m getting large, to a Gaussian distribution Y with $E(Y) = \sum_i E(W_i) \times r(x \succcurlyeq y)$ and $V(Y) = \sum_i V(W_i) \times |r(x \succcurlyeq y)|$. And the likelihood of *validation*, respectively *invalidation* of an "*at least as good as*" situation, denoted $lh(x \succcurlyeq y)$, may be assessed as follows:

$$lh(x \succcurlyeq y) \; = \; \begin{cases} 1.0 - P(Y \leqslant 0.0) & \text{if } E[\tilde{r}(x \succcurlyeq y)] > 0, \\ P(Y \leqslant 0.0) & \text{otherwise.} \end{cases} \qquad (3)$$

Example 2.1. Let us consider two decision alternatives x and y being evaluated on a family of 7 equi-significant criteria, such that four out of the seven criteria positively support that x outranks y, and three criteria support that x does not outrank y. In this case, $\tilde{r}(x \succcurlyeq y) = 4w - 3w = w$ where $W_i = w$ for $i = 1, ..., 7$ and the outranking situation is positively validated. Suppose now that the significance weights W_i appear only more or less equivalent and let us model this numerical uncertainty with independent triangular laws: $W_i \sim \mathcal{T}r(0, 2w, w)$ for $i = 1, ...7$. The expected credibility of the outranking situation, $E(\tilde{r}(x \succcurlyeq y)) = 4w - 3w = w$, will remain the same, however with a variance of $7 \times \frac{1}{6}w^2$. If we take a unit weight $w = 1$, we hence obtain a standard deviation of 1.08. Applying the CLT we notice that, under the given hypotheses, the likelihood $lh(x \succcurlyeq y)$ of obtaining a positive majority margin will be about $1.00 - P(\frac{\tilde{r}-1}{1.08} \leqslant 0.0) \approx 83\%$. A Monte Carlo simulation with 10 000 runs empirically confirms the effective convergence to a Gaussian: $\tilde{r}(x \succcurlyeq y) \rightsquigarrow \mathcal{N}(1.03, 1, 089)$ (see Figure 3), with an empirical probability of observing a negative majority margin $P(\tilde{r}(x \succcurlyeq y) \leqslant 0.0)$ of indeed about 17%.

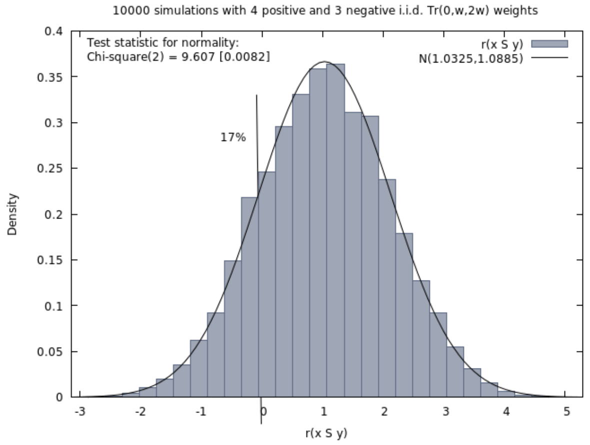

Figure 3. Distribution of outranking credibility $\tilde{r}(x \succcurlyeq y)$

Example 2.2. The second example concerns two decision alternatives a_1 and a_2 that are evaluated on a family of 7 criteria, denoted g_i of unequal significance weights w_i for $i = 1, ..., 7$ (see Tab. 1). The performances on the seven criteria are measured on a rational scale from 0 (worst) to 100 points (best). Let us suppose that both decision alternatives are evaluated as shown in Tab. 1. A performance difference of 10 points or less is considered insignificant, whereas a difference of 20 points and more is considered to be significant.

Table 1. Pairwise comparison of two decision alternatives

g_i / w_i	g_1 / 7	g_2 / 8	g_3 / 3	g_4 / 10	g_5 / 1	g_6 / 9	g_7 / 7
a_1	14.1	71.4	87.9	38.7	26.5	93.0	37.2
a_2	64.0	87.5	67.0	82.2	80.8	80.8	10.6
$a_1 - a_2$	-49.9	-16.1	+20.9	-43.5	-54.3	+12.2	26.5
$r(\succcurlyeq_i)$	-1	0	+1	-1	-1	+1	+1

The overall deterministic outranking credibility $r(a_1 \succcurlyeq a_2)$ (see [4]) is given as follows:

$$r(a_1 \succcurlyeq a_2) = \sum_{i=1}^{7} r(a_1 \succcurlyeq_i a_2) \times w_i \quad (4)$$
$$= -7 + 0 + 3 - 10 - 1 + 9 + 7 = +1 \quad (5)$$

The outranking situation "$(a_1 \succcurlyeq a_2)$" is thus positively validated (see Eq. 5). However, in case the given criteria significance weights (see Tab. 1) are not known with certainty, how confident can we be about the actual positiveness of $\tilde{r}(a_1 \succcurlyeq a_2)$? If we suppose now that the random significance weights W_i are in fact independently following a triangular continuous law on the respective ranges 0 to $2w_i$, the CLT approximation will make $\tilde{r}(a_1 \succcurlyeq a_2)$ tend to a Gaussian distribution with mean equal to $E(\tilde{r}(x \succcurlyeq y)) = +1$ and standard deviation equal to $\sqrt{\sum_i 1/6E(W_i)^2} = 6.94$. The likelihood of $\tilde{r}(a_1 \succcurlyeq a_2) > 0.0$ equals thus approximately $1.0 - P(\frac{z-1}{6.94} \leqslant 0.0) = 1.0 - 0.443 \approx 55.7\%$, a result we can again empirically verify with a Monte Carlo sampling of 10 000 runs (see Fig. 4). Under the given modelling of the

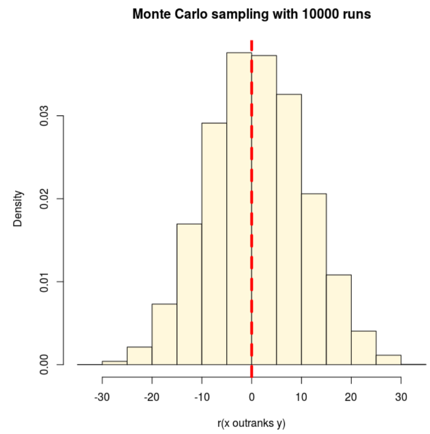

Figure 4. Distribution of outranking credibility $\tilde{r}(a_1 \succcurlyeq a_2)$

uncertainty in the setting of the criteria significance weights, the credibility of the outranking situation between alternatives a_1 and a_2 is neither convincingly positive, nor negative. The given relational situation may, hence, neither confidently be validated, nor, confidently invalidated.

3 Confidence level of outranking situations

Following the classic outranking difinition (see Roy [1], Bisdorff [4]), we may say from an epistemic point of view, that decision action x *outranks* decision action y, denoted $x \gtrsim y$), if

1. a *confident majority* of criteria **validates** a global outranking situation between x and y, and
2. *no considerably less performing* is observed on a discordant criterion.

Dually, decision action x *does not outrank* decision action y, denoted $(x \not\gtrsim y)$, if

1. *a confident majority* of criteria **invalidates** a global outranking situation between x and y, and
2. *no considerably better performing* situation is observed on a concordant criterion.

On a criterion i, we characterize a *considerably less performing* situation, called *veto* and denoted $\lll_i$, as follows:

$$r(x \lll_i y) = \begin{cases} +1 & \text{if} \quad x_i + v_i \leqslant y_i \\ -1 & \text{if} \quad x_i - v_i \geqslant y_i \\ 0 & \text{otherwise.} \end{cases} \quad . \quad (6)$$

where v_i, with $M_i \geqslant v_i > pr_i$, represents a lower-closed veto discrimination threshold. A corresponding dual *considerably better performing* situation, called *counter-veto* and denoted $\ggg_i$, is similarly characterized as:

$$r(x \ggg_i y) = \begin{cases} +1 & \text{if} \quad x_i - v_i \geqslant y_i \\ -1 & \text{if} \quad x_i + v_i \leqslant y_i \\ 0 & \text{otherwise.} \end{cases} \quad . \tag{7}$$

A global *veto*, or *counter-veto* situation is now defined as follows:

$$r(x \lll y) = \bigotimes_{i \in F} r(x \lll_i y) \tag{8}$$

$$r(x \ggg y) = \bigotimes_{j \in F} r(x \ggg_j y) \tag{9}$$

where $\otimes$ represents the epistemic polarising ([9]) or symmetric maximum ([10]) operator:

$$r \otimes r' = \begin{cases} \max(r, r') & \text{if} \quad r \geqslant 0 \wedge r' \geqslant 0, \\ \min(r, r') & \text{if} \quad r \leqslant 0 \wedge r' \leqslant 0, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

We observe the following semantics:

1. $r(x \lll y) = 1$ iff there exists a criterion i such that $r(x \lll_i y) = 1$ and there does not exist otherwise any criterion $j \in F$ such that $r(x \ggg_j y) = 1$.
2. Conversely, $r(x \ggg y) = 1$ iff there exists a criterion i such that $r(x \ggg_i y) = 1$ and there does not exist otherwise any criterion j such that $r(x \lll_j y) = 1$.
3. $r(x \ggg y) = 0$ if either we observe no very large performance differences or we observe at the same time, both a very large positive and a very large negative performance difference.

It is worthwhile noticing that $r(\not\lll)^{-1}$ is identical to $r(\ggg)$, both $\lll$ and $\ggg$ being, by construction, codual relations one to another.

The deterministic outranking characteristic $r(\succsim)$ may hence be defined as follows:

$$r(x \succsim y) = r(x \succcurlyeq y) \otimes_{i \in F} \big[-r(x \lll_i y) \big] \tag{11}$$

And in particular,

1. $r(x \succsim y) = r(x \succcurlyeq y)$ if no very large positive or negative performance differences are observed,
2. $r(x \succsim y) = 0$ if a veto and a counter-veto situation are conjointly occurring;
3. $r(x \succsim y) = 1$ if $r(x \succcurlyeq y) \geqslant 0$ and $r(x \ggg y) = 1$,
4. $r(x \succsim y) = -1$ if $r(x \succcurlyeq y) \leqslant 0$ and $r(x \lll y) = 1$.

When considering now the criteria significance weights to be random variates, $r(x \succsim y)$ becomes a random variable via the random characteristic $\tilde{r}(x \succcurlyeq y)$.

$$\tilde{r}(x \succsim y) = \tilde{r}(x \succcurlyeq y) \otimes_{i \in F} \big[-r(x \lll_i y) \big] \tag{12}$$

In case 1. we are back to the unpolarised "*at least as good as*" situation discussed in the previous section. In case 2., the resulting constant indeterminate outranking characteristic value 0 is in fact independent of any criterion significance. Only cases 3. and 4. are of interest here. If $E(\tilde{r}(x \succcurlyeq y)) \geqslant 0$, we are in case 3. where strictly negative characteristics will be given the indeterminate characteristic 0, and the others, a polarised +1 value. Similarly, if $E(\tilde{r}(x \succcurlyeq y)) \leqslant 0$ we are in case 4., strictly positive characteristics $r(x \succcurlyeq y) > 0$ will be given the indeterminate value 0, and the others, the polarised −1 value.

By requiring now a certain level α of likelihood for effectively validating all pairwise outranking situations, we may thus enforce the actual confidence we may have in the valued outranking digraph. For any outranking situation $(x \succsim y)$ we obtain:

$$\hat{r}_\alpha(x \succsim y) = \begin{cases} E\big[\tilde{r}(x \succsim y)\big] & \text{if } lh(x \succcurlyeq y) \geqslant \alpha, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

If, for instance, we would require that an outranking situation $(x \succsim y)$, to be validated, respectively invalidated, must admit a likelihood $lh(x \succcurlyeq y)$ of $\alpha = 90\%$ or more, any positively or negatively polarising of the "at least as good as" statement will only occur in case of sufficient likelihood. Noticing that $E\big[\tilde{r}(x \succsim y)\big] = r(x \succsim y)$, we safely preserve, hence, in our stochastic modelling, all the nice structural properties of the deterministic outranking relation (see Eq. 11), like *weak completeness* and *coduality*, that is the dual of the outranking relation ($\not\succsim$) corresponds to the asymmetric part ($\precnsim$) of its converse relation (see Bisdorff [4]).

Example 3.1. We may illustrate our uncertainty modelling approach with a small random performance tableau (see Tab. 2) showing the evaluations of seven decision alternatives on the same family of performance criteria we used for Example 2.2. To operate with a full fledged outranking model, we furthermore consider that a performance difference of 80 points and more will trigger a veto or counter-veto situation (see [4]).

Table 2. Random performance tableau

g_i	w_i	a_1	a_2	a_3	a_4	a_5	a_6	a_7
g_1	7	14.1	64.0	73.4	36.4	30.6	85.9	97.8
g_2	8	71.4	87.5	61.9	84.7	60.4	54.5	45.8
g_3	3	87.9	67.0	25.2	34.2	87.3	43.1	30.4
g_4	10	38.7	82.2	94.1	86.1	34.1	97.2	72.2
g_5	1	26.5	80.8	71.9	21.3	56.4	88.1	15.0
g_6	9	93.0	80.8	23.2	57.2	81.4	16.6	93.0
g_7	7	37.2	10.6	64.8	98.9	69.9	24.7	13.6

Thresholds: $ind_i = 10.0$, $pr_i = 20$, and $v_i = 80$ for $i \in F$.

When using the deterministic criteria significance weights shown in Tab. 2, we obtain the bipolarly valued outranking relation shown in Tab. 3. We recover there the weakly positive credibility ($r(a_1 \succsim a_2) = +1/45$) of the outranking situation between alternative a_1 and alternative a_2 discussed in Example 2.2. Notice also the slightly negative credibility (−5/45) of the outranking situation between alternative a_1 and a_3. Notice, furthermore the veto and counter-veto situations we observe when comparing alternatives a_1 and a_7, a_2 and a_4, as well as, a_4 and a_7. How confident are all these pairwise preferential situations when the significance weights are not precisely given? Assuming that the criteria significance weights w_i are in fact random variates distributed following independent triangular laws $\mathcal{T}(0, 2w_i, w_i)$ for $i = 1, ..., 7$, we obtain

Table 3. Deterministic credibility of $(x \succsim y)$

$r(\succsim) \times 45$	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	**+1**	**−5**	−11	+22	+9	**0**
a_2	+16	-	+21	**0**	+25	+14	+22
a_3	+21	+5	-	−3	+21	+34	+13
a_4	+21	**+45**	+29	-	+19	+19	**+45**
a_5	+28	−7	+10	−5	-	+9	+2
a_6	+6	+5	+31	−3	+7	-	+20
a_7	**+45**	+11	+1	**0**	+15	+13	-

Table 4. CLT likelihood of the $(x \succcurlyeq y)$ situations

lh	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	.56	.74	.94	1.0	.88	.92
a_2	.99	-	1.0	**1.0**	1.0	.99	1.0
a_3	1.0	.74	-	.65	1.0	1.0	.95
a_4	1.0	**.74**	1.0	-	.99	1.0	.95
a_5	1.0	.82	.90	.74	-	.88	.62
a_6	.83	.74	1.0	.65	.82	-	1.0
a_7	.85	.95	.56	.78	.98	.97	-

the CLT likelihoods shown Tab. 4. If we, now, require for each "*at least as good as*" situation $(x \succcurlyeq y)$ to admit a likelihood of 90% and more for convincingly validating, respectively invalidating, the corresponding outranking statement $(x \succsim y)$, we obtain the result shown in Tab. 5.

Table 5. 90% confident outranking characteristics (×45)

$\hat{r}_{90\%}(x \succsim y)$	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	0	0	−11	+22	0	0
a_2	+16	-	+21	0	+25	+14	+22
a_3	+21	0	-	0	+21	+34	+13
a_4	+21	0	+29	-	+19	+19	+45
a_5	+28	0	+10	0	-	0	0
a_6	0	0	+31	0	0	-	+20
a_7	0	+11	0	0	+15	+13	-

We notice there that, for instance, the outranking situations $(a_1 \succsim a_2)$ and $(a_1 \not\succsim a_3)$, with likelihoods 56%, resp. 73% – lower than 90%– are both put to doubt. Similarly, the +45 polarised outranking situation $(a_7 \succsim a_1)$ appears not confident enough. Same happens to +45 polarised situation $(a_4 \succsim a_2)$. Whereas situation $(a_4 \succsim a_7)$ remains confidently polarised to +1. In total 16 pairwise outranking statements, out of the potential 7×6 statements, are thus considered not confident enough. At required confidence level of 90%, their credibility $\hat{r}_{90\%}(x \succsim y)$ is put to the indeterminate value 0. It is worthwhile noticing that all outranking situations, showing a majority margin between $\pm 9/45$ (between 40 and 60%) are thus not confident enough and consequently put to doubt (characteristic value 0).

4 Exploiting the confident outranking digraph

Many MCDA decision aiding problematiques like best choice, ranking, sorting, or clustering recommendations based on pairwise outranking situations, rely on majority cuts of the corresponding valued outranking digraph (see [11, 12, 13]).

Example 4.1. The previous example 3.1 gives the hint how we may appreciate the very confidence we may have in a given majority when the criteria significance weights are not precisely given. We may, for instance, notice that alternative a_4 gives apparently the only Condorcet winner in the deterministic outranking digraph and will hence be recommended in the RUBIS decision aid approach as best choice (see [13]). In the 90% confident outranking digraph, however, alternatives a_2 and a_4 both give two equivalent weak Condorcet winners, and may, hence, be both recommended as potential RUBIS best choice candidates; a recommendation more convincingly supported than the deterministic one, when considering in fact the excellent performances of alternative a_2 compared to a_4 (see Tab. 6).

Table 6. Pairwise comparison of alternatives a_4 and a_2

g_i / w_i	g_1 7	g_2 8	g_3 3	g_4 10	g_5 1	g_6 9	g_7 7
a_4	36.5	84.7	34.2	86.1	21.3	57.2	98.9
a_2	60.0	87.5	67.0	82.2	80.8	80.8	10.6
$-$	-27.5	-2.8	-32.8	+3.8	-59.2	-23.6	+88.8
$\succcurlyeq_j$	−1	+1	−1	+1	−1	−1	+1
$\ggg_j$	0	0	0	0	0	0	+1

$[r(a_3 \succcurlyeq a_2) = +5 \,\wedge\, r(a_4 \ggg a_2) = +1] \Rightarrow r(a_4 \succsim a_2 = +45;$

$[lh_{.90}(a_4 \succcurlyeq a_2) \;=\; .74] \;\Rightarrow\; \hat{r}(a_4 \succsim a_2) \;=\; 0.$

Being confidently *at least a s good as* alternative a_4 ($lh(a_2 \succcurlyeq a_4) = 100\%$, see Tab. 4), alternative a_2 shows four excellent performances over 80.0, whereas alternative a_4 only shows three such high evaluations. The actual difference between the *deterministic* and the *confident* best choice recommendation stems in fact from the not confident enough polarisation of the counter-veto affecting the performance comparison between a_4 and a_2 ($lh(a_4 \succcurlyeq a_2) = 74\% \;<\; \alpha = 90\%$, see Tab. 4). Hence, alternative a_4 does no more appear alone as the Condorcet winner. Both, alternatives a_2 and a_4 appear as *confident weak Condorcet* winners, hence their joint recommendation as *confident best choice* candidates.

Knowing a priori the distribution of the significance weight of each criterion will genuinely be sufficient in practice for computing, with the so given means and variances, the CLT based likelihood of the fact that a bipolar outranking characteristics $r(x \succsim y)$ is positively validating, respectively negatively invalidating, the outranking situation "$(s \succsim y)$". The quality of the CLT convergence will, however, depend, first, on the number of effective criteria, i.e. non abstaining ones, involved in each pairwise comparison and, secondly, on the more or less differences in shape of the individual significance weight distributions. Therefore, with tiny performance tableaux, less than 25 decision actions and less than 10 criteria, we may estimate more precisely the actual likelihood of all pairwise outranking situations with a Monte Carlo (MC) simulation consisting of a given number of independent runs. Indeed, the present computational power available, even on modest personal computers, allow us to sufficiently sample a given outranking digraph construction.

Example 4.2. If we sample, for instance, 10 000 MC simulations of the previous outranking relation (see Tab. 3), by keeping the same uncertainty modelling of the criteria significances with random weights distributed like $\mathcal{T}(0, 2w_i, w_i)$, we obtain same empirical likelihoods (see Tab. 7).

Table 7. Empirical likelihoods of $(x \succcurlyeq y)$ with a MC sampling of 10 000 runs

p-value	a01	a02	a03	a04	a05	a06	a07
a01	-	.55	.74	.95	1.0	.88	.92
a02	.99	-	1.0	1.0	1.0	.99	1.0
a03	1.0	.74	-	.65	1.0	1.0	.96
a04	1.0	.75	1.0	-	1.0	1.0	.96
a05	1.0	.82	.90	.75	-	.88	.61
a06	.83	.74	1.0	.65	.82	-	1.0
a07	.85	.96	.55	1.0	.99	.97	-

We may thus verify again the very accurate convergence (in the order of $\pm 1\%$) of the CLT likelihoods, a convergence we already observed in Example 2.2, even with a small number of criteria.

Conclusion

In this paper we illustrate some simple models for tackling uncertain significance weights: uniform, triangular and beta laws. Applying the Central Limit Theorem, we are able to compute under these uncertainty models the actual likelihood of any pairwise *at least as good as* situations. This operational result, by adequately handling potential veto and counter-veto situations, allows to enforce a given confidence level on the corresponding outranking situations. On a small illustrative best choice problem, we eventually show the pragmatic decision aid benefit one may expect from exploiting a confident versus a classic deterministic outranking digraph. Acknowledging this operational benefit, one may finally be tempted to extend the uncertainty modelling, as in the SMAA approach, to the marginal performances. This is however, not needed, as traditionally the performance discrimination thresholds proposed in the outranking approach may well take care of any imprecision and uncertainty at this level.

References

[1] Roy B (1985) Méthodologie Multicritère d'Aide à la Décision. Economica Paris 1, 3

[2] Tervonen, T. and Figueira, J. R. (2008). A survey on stochastic multicriteria acceptability analysis methods. Journal of Multi-Criteria Decision Analysis (Wiley) 15(1-2): 1-14 1

[3] Bisdorff, R. (2000). Logical foundation of fuzzy preferential systems with application to the electre decision aid methods. Computers and Operations Research (Elsevier) 27:673-687 1

[4] Bisdorff, R. (2013) On polarizing outranking relations with large performance differences. Journal of Multi-Criteria Decision Analysis (Wiley) 20:3-12 1, 3, 4

[5] Bisdorff, R. (2002). Logical Foundation of Multicriteria Preference Aggregation. In: Bouyssou D et al (eds) Essay in Aiding Decisions with Multiple Criteria. Kluwer Academic Publishers 379-403 1

[6] Bisdorff, R., P. Meyer and Th. Veneziano (2009). Inverse analysis from a Condorcet robustness denotation of valued outranking relations. In F. Rossi and A. Tsoukiás (Eds.), Algorithmic Decision Theory. Springer-Verlag Berlin Heidelberg, LNAI 5783: 180-191 1

[7] Bisdorff, R., P. Meyer and Th. Veneziano (2014). Elicitation of criteria weights maximising the stability of pairwise outranking statements. Journal of Multi-Criteria Decision Analysis (Wiley) 21: 113-124 1

[8] Dias, L. C. (2002). Exploring the Consequences of Imprecise Information in Choice Problems Using ELECTRE. In Bouyssou, D. et al. (eds) Aiding Decisions with Multiple Criteria (Springer ORMS) 44: 175-193 1

[9] Bisdorff, R. (1997). On computing fuzzy kernels from l-valued simple graphs. In 5th European Congress on Intelligent Techniques and Soft Computing EUFIT'97, 1: 97- 103 4

[10] Grabisch M., J.-L Marichal, R. Mesiar, and E. Pap (2009). Aggregation functions. Encyclopedia of Mathematics and its Application. Cambridge University Press 4

[11] Bisdorff, R. (2002), Electre like clustering from a pairwise fuzzy proximity index. European Journal of Operational Research EJOR (Elsevier) 138/2: 320-331 5

[12] Bisdorff, R., M. Pirlot and M. Roubens (2006). Choices and kernels from bipolar valued digraphs. European Journal of Operational Research (Elsevier) 175: 155-170 5

[13] Bisdorff, R., P. Meyer and M. Roubens (2008). RUBIS: a bipolar-valued outranking method for the choice problem. 4OR - A Quarterly Journal of Operations Research (Springer-Verlag) 6 (2): 143-165 5