HAL
open science

Analyse de la signifiance de diverses procédures d'agrégation multicritère (décembre 2002)

Bernard Roy, Jean-Marc Martel

To cite this version:

Bernard Roy, Jean-Marc Martel. Analyse de la signifiance de diverses procédures d'agrégation multicritère (décembre 2002). 2005. hal-00004107

HAL Id: hal-00004107

https://hal.science/hal-00004107v1

Preprint submitted on 1 Feb 2005

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Table des matières

Abstract i

Résumé ii

1. Introduction 1

2. La théorie du mesurage : Un bref aperçu 2
   2.1. Généralités 2
   2.2. Typologie des échelles de mesure 5

3. La notion de signifiance 9
   3.1. Définition 9
   3.2. Illustration du concept de signifiance à partir de la somme pondérée 11

4. Analyse de quelques PAMC 14
   4.1. PAMC de type critère unique de synthèse 14
      4.1.1. Généralités 14
      4.1.2. MAUT (Keeney et Raiffa, 1976) 15
      4.1.3. SMART (Edwards, 1977 ; Von Winterfield et Edwards, 1986) 17
      4.1.4. TOPSIS (Hwang et Yoon, 1981) 19
   4.2. PAMC de type surclassement de synthèse 20
      4.2.1. Généralités 20
      4.2.2. ELECTRE I 21
      4.2.3. ELECTRE IS et ELECTRE III 23
      4.2.4. PROMÉTHÉE I et II 29

5. Conclusion 31

Références 33

Analysis of meaningfulness of multicriterion aggregation procedures

Abstract

The meaningfulness concept (see Suppes, 1954) is issue of the measurement theory. This theory (see Krantz et al., 1971) concerns the way in which we can represent specific information (in particular qualitative nature) related to a given class of phenomena, by a set of numerical values so that this representation reflects suitably the specific properties of the concerned phenomena. To have a meaningful statement (relating to the class of phenomena) founded on the calculation which uses numerical representation, it's necessary that its validity or invalidity must not be affected when measurement is changed, provided that it remains acceptable.

In decision-making aid, a criterion assigns a performance to each potential action (level which ca be always numerical) on a specific scale. This representation aims to take into account a preference system related to a given point of view. When we consider several points of view, the aim of the multicriterion aggregation procedure (MCAP) is to build a global preference system which takes into account all the points of view. The ensuing statement of type « a is strictly preferred to b » are meaningful only if they remain unchanged under admissible transformation of scale for any point of view.

In this work, we analyse the meaningfulness of statements resulting from a specific number of MCAP usually used in decision-making aid (section 3 and 4). We present as preliminary the elements of the measurement theory used as support to lead this analysis. Indeed, this analysis plays a central role in a traditional typology of the scales of measurement. This led us (§ 2.2) to propose an adapted definition for our matter.

Keywords : meaningfulness concept, measurement theory, multicriterion aggregation procedure.

Analyse de signifiance de diverses procédures d'agrégation multicritère

Résumé

La notion de signifiance (voir Suppes, 1954) est issue de la théorie du mesurage. Cette théorie (voir Krantz et al., 1971) concerne la manière dont on peut représenter certaines informations (notamment de nature qualitative) ayant trait à une classe de phénomènes donnés, par un ensemble de valeurs numériques et cela de telle sorte que cette représentation reflète convenablement certaines propriétés des phénomènes concernés. Pour qu'une assertion (relative à la classe de phénomènes) fondée sur un calcul faisant intervenir la représentation numérique soit signifiante, il est nécessaire que sa validité ou non-validité ne soit pas affectée lorsqu'on change de mesurage, pourvu qu'il reste convenable.

En aide à la décision, un critère conduit à attribuer aux actions potentielles une performance (échelon que l'on peut toujours regarder comme numérique) sur une certaine échelle. Cette représentation a pour objet de prendre en compte un système de préférences lié à un point de vue donné. Lorsque l'on fait intervenir plusieurs points de vue, la procédure d'agrégation multicritère (PAMC) a pour objet de bâtir un système de préférences global prenant en compte tous les points de vue. Les assertions qui en découlent du type «*a* est strictement préféré à *b*» ne sont signifiantes que si elles sont invariantes lorsque l'on passe d'un mesurage convenable à un autre pour un point de vue donné quel qu'il soit.

Dans cet article, on analyse la signifiance des assertions auxquelles conduisent un certain nombre de PAMC couramment utilisées en aide à la décision (sections 3 et 4). On présente au préalable les éléments de la théorie du mesurage sur lesquels on a pris appui pour conduire cette analyse. Cette dernière fait jouer un rôle central à une typologie classique des échelles de mesure. Ceci nous a amenés (§ 2.2) à en proposer une définition adaptée à notre propos.

Mots clés : Notion de signifiance, théorie de mesurage, procédure d'agrégation multicritère.

1. Introduction

D'une façon générale, les valeurs numériques apparaissent presque à tous les niveaux d'un raisonnement ou d'une analyse scientifique et l'aide à la décision n'y échappe pas. L'introduction de ces valeurs numériques est souvent nécessaire pour prendre en compte des concepts ou des observations empiriques. Ces recours à l'usage du nombre constituent sans doute une étape très importante dans la formulation (ou modélisation) d'un problème. D'ailleurs, cette importance est largement reconnue et même envisagée dans le domaine des sciences sociales, en particulier dans l'aide à la décision. Dans ce domaine, le passage des mots, des observations ou de l'information qualitative aux valeurs numériques n'a été accompli, du point de vue théorique, que partiellement. En effet, on peut facilement remarquer en examinant les différents cas étudiés dans le domaine de l'aide à la décision, notamment dans l'aide multicritère à la décision, l'absence d'une méthodologie générale et cohérente permettant le passage de l'information qualitative aux valeurs numériques.

Si l'on considère qu'une décision est basée sur les conséquences (appréhendées à partir de critères multiples) des actions potentielles, ces conséquences doivent être évaluées et, pour cela, le processus d'aide à la décision incorpore généralement pour son fonctionnement une procédure d'attribution de valeurs numériques. De plus, il est évident que le traitement de ces conséquences est une étape essentielle dans toute méthode multicritère. L'attribution de valeurs numériques facilite ce traitement (l'agrégation) lorsque ce dernier est requis.

Cependant, parallèlement, il existe une théorie, la théorie du mesurage, qui concerne la représentation de structures empiriques par des valeurs numériques. Toutefois, cette théorie n'a pas encore été suffisamment explorée dans le domaine de l'aide à la décision comparativement aux domaines de la psychologie et de l'économie par exemple. Elle essaie d'établir les bases, en termes d'axiomatisation, d'une telle représentation.

La théorie de mesurage traite de la représentation d'un système empirique sous-jacent supposé préexister. Celui-ci n'a, dans bien des cas, son véritable équivalent lorsqu'il s'agit d'asseoir la construction d'un ou plusieurs critères dont on a besoin pour pendre en compte tel ou tel point de vue en aide à la décision. C'est pourquoi, bien souvent, les critères ne sont pas conçus pour être des représentations fidèles d'un système de préférence empirique préexistant (voir par exemple Bouyssou et al., 2000, chapitre 4). Il n'en demeure pas moins que la théorie du mesurage peut contribuer (comme on espère le mettre en évidence dans la suite de cet article) à s'interroger, de façon très pertinente pour la prise en compte des points de vue, sur la signification qu'il est nécessaire de reconnaître aux nombres manipulés

pour que la façon dont ceux-ci sont traités dans les calculs puisse conduire à des conclusions qui ne soient pas trop discutables. En absence d'un système empirique parfaitement défini qu'il s'agit de décrire, on peut en effet s'interroger sur cette signification dès l'instant où l'on adopte une démarche constructiviste (cf. Roy, 1993) qui amène à regarder les critères non pas comme une description fidèle d'une réalité parfaitement objective mais comme traduisant une hypothèse de travail adaptée à l'objectif de l'aide et à la nature des phénomènes qu'il s'agit de prendre en compte. Chaque fois que l'on utilise le nombre pour exploiter des informations qui sont de nature plus qualitative que quantitative, le risque est grand de tomber dans un arithmomorphisme dénué de sens. La théorie du mesurage nous semble pouvoir apporter des garde-fous pour se protéger contre ce risque. Dans cet article, on analyse la signifiance de diverses procédures d'agrégation multicritère (PAMC) en mettant l'accent sur le traitement de l'information selon trois aspects : la représentation, l'unicité et la signifiance.

Nous présentons dans la deuxième section un bref aperçu de la théorie du mesurage ainsi que la définition et la description des principales échelles de mesure et de leurs caractéristiques. Nous présentons une classification des échelles de mesure et abordons les notions d'échelles régulières et de transformations admissibles.

Dans la troisième section, nous exposons la notion de signifiance que nous illustrons à partir de la somme pondérée. Dans la section suivante, une analyse du point de vue de la signifiance est faite pour les PAMC suivantes : MAUT, SMART, TOPSIS, ELECTRE I, IS et III, PROMETHEE I et II. Finalement, nous concluons en résumant les principaux résultats de cette étude.

2. La théorie du mesurage : Un bref aperçu

2.1. Généralités

Plusieurs définitions dans la littérature ont été données au terme mesurage. Ainsi, Campbell (1938, p. 126) définit le terme mesurage comme étant *: "The assignment of numerals to represent properties of material systems other than numbers, in virtue of the laws governing these properties*".

Pour Russel, mesurage signifie : "*Measurement of magnitudes is, in its most general sense, any method by which a unique and reciprocal correspondence is established between all or some of the magnitudes of a kind and all or some the numbers, integral, rational, or real as the case may be*" (1938, p. 176).

Pour Stevens, mesurage est : "*Measurement is the assignment of the numerals to objects or events according to rules*" (1951b, p. 22).

Ces trois définitions constituent les principaux points de vue que l'on peut trouver dans la littérature sur la théorie du mesurage. Toute autre définition peut, à quelques modifications près, être représentée par l'un des trois points de vue précédents.

Ces définitions du mesurage suggèrent que les nombres[1] associés doivent correspondre à, ou représenter, ou préserver certaines relations observées. Ainsi, le mesurage d'un ensemble d'observations empiriques (d'informations qualitatives) consiste à construire une fonction f à valeurs réelles définies sur ce dernier. De plus, cette fonction doit avoir la particularité de conserver toutes [2] les propriétés de l'ensemble d'observations empiriques.

Par exemple, supposons pour le mesurage de la masse que nous ayons l'information suivante : "l'objet a est plus lourd que l'objet b". Nous pouvons représenter cette information en attribuant à "l'objet a" une valeur numérique plus grande que celle attribuée à "l'objet b". Autrement dit, nous pouvons représenter cette information par la fonction suivante :

$$\text{«L'objet a est plus lourd que l'objet b»} \Leftrightarrow f(a) > f(b).$$

De même, en aide à la décision, si nous jugeons qu'une action "a est préférée à une action b" selon un point de vue, nous devons associer à a une valeur numérique g(a) plus grande que celle g(b) associée à b et ainsi :

$$\text{«a } P_g \text{ b»} \Leftrightarrow g(a) > g(b),$$

Le nombre réel g(a) est appelé le mesurage (la performance) de a et l'application $g : A \rightarrow 3$ (où A est l'ensemble des actions potentielles et 3 l'ensemble des réels) constitue ce qu'on appelle un critère.

Revenons à l'exemple de la fonction f introduite pour mesurer la masse. On peut vouloir prendre en compte la propriété suivante: «La masse de la "concaténation" de deux objets a et b est la somme de leur masse». Dans ce cas, la fonction f devra satisfaire une deuxième condition qui représentera la propriété de concaténation $f(a \oplus b) = f(a) + f(b)$. Cet exemple illustre ce que l'on attend du mesurage : le mesurage d'un ensemble d'observations.

[1] Le mesurage avec des nombres n'est pas la seule forme de mesurage, mais c'est certainement la plus usuelle, notamment en aide à la décision.
[2] du moins celles jugées pertinentes dans le contexte considéré.

En résumé, le mesurage d'un ensemble d'observations empiriques consiste à construire un homomorphisme de ce système relationnel empirique vers un système relationnel numérique. Ainsi, dans l'exemple de la masse, si la relation $a \succ b$ signifie que l'objet a est plus lourd que l'objet b, si l'on désigne la concaténation par $\oplus$ et si A est l'ensemble des objets, alors la fonction f est un homomorphisme de M_1 dans M_2 où $M_1 = (A, \oplus, \succ)$ est le système relationnel empirique et $M_2 = (3, +, >)$ est le système relationnel numérique.

La théorie du mesurage cherche principalement à répondre aux questions suivantes :

- Existe-t-il un système relationnel numérique permettant de représenter le système relationnel empirique ?
- Cette représentation numérique est-elle unique ?
- Quels traitements mathématiques sont autorisés (sont signifiants) avec cette représentation numérique ? Les énoncés (recommandations) issus de ces traitements restent-ils les mêmes s'ils sont effectués avec une autre représentation numérique admissible (cas de non unicité), c'est-à-dire représentant elle aussi le système relationnel empirique ?

Cette dernière question, qui est centrale dans cette étude, sera l'objet de la section suivante (section 3).

La question de la représentation fait intervenir un ensemble de conditions (axiomes) et de théorèmes (Roberts, 1979) concernant l'existence d'un homomorphisme du système relationnel empirique dans un système relationnel numérique. A titre d'exemple, dans le contexte de l'aide à la décision, on peut mentionner des axiomes fréquemment évoqués comme ceux de :

La transitivité :
si, pour tout a, b et c $\in$ A, on observe a P_g b et b P_g c, alors on devrait avoir a P_g c,

L'asymétrie :
si, pour tout a, b $\in$ A, on observe a P_g b, alors on devrait avoir non b P_g a,

La transitivité négative :
si, pour tout a, b et c $\in$ A, on observe non a P_g b et non b P_g c, alors on devrait avoir non a P_g c.

ou encore le **théorème de Cantor** : Soit A un ensemble fini ou dénombrable sur lequel est définie une relation P (traduisant par exemple une préférence). Il existe une fonction à valeurs réelles V définie sur A telle que :

$$\text{pour tout } a, b \in A,\ a\ P\ b \Leftrightarrow V(a) > V(b)$$

si et seulement si P est asymétrique et respecte la transitivité négative.

Une fois que l'on est parvenu à établir cet homomorphisme, c'est-à-dire cette représentation numérique, se pose alors la question de son unicité. En général, cette représentation numérique n'est pas unique. Selon Barzilai (1998), la forme typique d'un théorème d'unicité est la suivante : si f est un homomorphisme d'un système relationnel M_1 dans un système relationnel M_2, alors n'importe quel autre homomorphisme h de M_1 dans M_2 est donné par :

$$h = \phi\ (f)\ (\text{ou } \phi \circ f)\ ^{3}$$

où ϕ appartient à un ensemble spécifié Φ de transformations ne dépendant pas de f.

Dans notre exemple de la masse, il est clair que si Φ est l'ensemble des transformations $\Phi(a) = cf(a)$ où c est une constante positive, alors si f est une représentation numérique appropriée pour associer la masse à l'objet a, il en est de même pour la représentation numérique $h = \phi(f)$ (ce qui correspond à un changement d'unité).

On se place dans cet article dans le contexte des **échelles régulières** (cf. § 2.2) selon Roberts (1979, 1994). Cela signifie que s'il existe deux homomorphismes f et h d'un système relationnel dans un autre système relationnel, alors il existe une transformation admissible ϕ sur f(A) telle que $h = \phi \circ f$ où A est l'ensemble sous-jacent au premier système relationnel. On ne considère donc que les cas que Barzilai (1998) qualifie de non pathologiques, c'est-à-dire que si f et h sont deux homomorphismes de M_1 dans M_2, alors :

$$f(a) = f(b) \Leftrightarrow h(a) = h(b).$$

Cette question est étroitement reliée à celle de la signifiance et sera abordée ici à travers une typologie des échelles de mesure. Pour définir cette typologie, nous supposons enfin l'ensemble A fini ou dénombrable.

2.2. Typologie des échelles de mesure

Dans le reste de cet article, on appellera :

[3] Les deux notations sont utilisées.

- **représentation numérique** (en abrégé rn) de A : une fonction f utilisée pour associer à tout objet a de A le nombre f(a) censé représenter une certaine information ;

- **mesurage** ou **représentation numérique appropriée** (rna) de A vis-à-vis d'une propriété donnée de l'information : une rn f qui applique A dans un ensemble de nombres de telle sorte que certaines des caractéristiques des nombres que constituent les codages f(a) ait un sens concret lié à la propriété considérée de l'information ;

- **échelle** (ou **échelle de mesure**) : le couple formé par (i) un ensemble de nombres susceptibles d'être utilisés pour coder une information relative aux objets de A et (ii) un mesurage (rna) appliquant A dans cet ensemble de nombres ; ces nombres sont appelés échelons de l'échelle.

Parmi les différents types d'échelles et leur classification exposés dans la littérature, nous retenons ceux développés par Stevens (1946) et adoptés par Roberts (1979).

Considérons une série de valeurs numériques. Elle peut être pertinente pour attribuer une signification à l'une des trois caractéristiques suivantes :

- **Caractéristique d'ordre** : rangement par valeurs croissantes ou décroissantes des nombres constituant la série.

- **Caractéristique de distance** : rangement par valeurs croissantes ou décroissantes des différences entre tous les couples de nombres constituant la série.

- **Caractéristique d'origine** : rôle joué par la valeur 0 vis-à-vis des nombres constituant la série.

Ces caractéristiques peuvent servir de point de référence pour la construction d'une échelle. En effet, face à un ensemble d'observations empiriques (d'informations qualitatives), nous essaierons de mettre en évidence les caractéristiques de l'ensemble empirique qui seront analogues à celles d'un ensemble de valeurs numériques.

Selon la signification que l'on peut accorder aux caractéristiques ci-dessus, on peut distinguer, avec Stevens (1946) et Roberts (1979), cinq types d'échelles :

Échelle nominale : Échelle faisant intervenir un ensemble de nombres qui ne donnent sens à aucune des trois caractéristiques ci-dessus car le mesurage [4] a seulement pour fonction de repérer et nommer les objets de A : le nombre qui leur est ainsi associé ne sert qu'à les désigner et à les différencier. Il s'ensuit que toute transformation injective de ces nombres sur l'ensemble A définit une représentation numérique appropriée avec l'ensemble des nombres considérés.

Échelle ordinale : Échelle faisant intervenir un ensemble de nombres qui donnent sens à la caractéristique d'ordre car tous les objets de A doivent être rangés exactement de la même façon dans n'importe quelle représentation numérique appropriée ; l'échelle sera dite **seulement ordinale** si, dès l'instant où f est une rna et ϕ une transformation monotone strictement croissante, alors $h = \phi \circ f$ est une rna et toutes les rna sont déduites de f de cette façon.

Échelle intervalle : Échelle ordinale à propos de laquelle on peut donner sens à la caractéristique de distance car le rapport des différences des nombres associés à deux couples d'objets distincts doit avoir la même valeur dans n'importe quelle représentation numérique appropriée ; il s'ensuit que si f est une rna, toutes les rna sont de la forme $h = \alpha f + \beta$ avec $\alpha > 0$.

Échelle ratio : Échelle intervalle à propos de laquelle on peut donner sens à la caractéristique d'origine car le rapport des nombres associés à deux objets distincts doit avoir la même valeur dans n'importe quelle représentation numérique appropriée ; il s'ensuit que si f est une rna, toutes les rna sont de la forme $h = \alpha f$ avec $\alpha > 0$.

Échelle absolue : Échelle ratio qui n'autorise qu'une seule représentation numérique appropriée.

L'échelle ratio devient une échelle absolue dès l'instant où l'on impose le choix de l'unité et l'échelle intervalle devient une échelle ratio dès l'instant où l'on fixe l'origine. L'échelle ratio est une échelle qui **mesure une quantité** : d'une part elle donne sens à l'absence de quantité, d'autre part toute mesure peut s'interpréter comme l'addition d'un nombre entier ou fractionnaire de l'unité de la quantité considéré.

[4] Dans ces conditions, l'emploi du terme «mesurage» peut sembler inapproprié, de même que celui d'échelle.

Faisons observer que la façon de donner sens à la caractéristique de distance dans une échelle intervalle est extrêmement forte. En aide à la décision, on est souvent en présence d'un type d'échelle intermédiaire entre l'échelle seulement ordinale et l'échelle intervalle (au sens fort introduit ci-dessus) qui est définie comme suit :

Échelle intervalle faible[5] : Échelle ordinale à propos de laquelle on peut donner sens à la caractéristique de distance car le rangement (et non pas le rapport) des différences des nombres associés à deux couples d'objets distincts doit être exactement le même dans n'importe quelle représentation numérique appropriée.

Pour ce type d'échelle, il n'existe pas (comme pour les types précédents) de caractérisation simple de l'ensemble Φ des transformations ϕ qui permette de passer d'une rna à n'importe quelle autre.

On remarquera que les types d'échelles qui viennent d'être définis possèdent la propriété de régularité énoncée à la fin du 2.1. Considérons en effet deux objets a et b de A qui reçoivent des mesures distinctes $f(a) \neq f(b)$ avec une rna f. Avec toute autre rna h, on aura nécessairement $h(a) \neq h(b)$, aussi bien pour une échelle nominale que pour une échelle ordinale.

Faisons enfin observer que l'ensemble Φ des transformations dites admissibles qui font passer d'une rna f à une autre rna ne dépend pas de f pour les types d'échelles considérés ici. De plus, ces ensembles Φ forment une suite emboîtée (cf. tableau 1) allant du plus restreint avec l'échelle absolue (Φ = la seule transformation identité) jusqu'au plus riche avec l'échelle nominale (Φ = ensemble des transformations injectives). Ces types d'échelles constituent de ce fait une hiérarchie qui permet de parler de **niveau de mesurage** : on dira par exemple que l'échelle ratio constitue un niveau de mesurage plus élevé que l'échelle intervalle, laquelle est elle-même plus élevée que l'échelle seulement ordinale.

[5] Ce nouveau type d'échelle est connu en anglais sous la dénomination *ordered metric scale* ; la terminologie adoptée ici vise à souligner la proximité qui existe entre une échelle de type intervalle (au sens le plus usuel) et ce nouveau type d'échelle : dans l'une comme dans l'autre, la comparaison d'intervalles doit être chargée de signification. Toutefois, celle-ci est plus forte dans le premier de ces deux types (d'où le qualificatif fort que l'on peut lui accoler pour éviter toute confusion) que dans le second.

Tableau 1 : Classification des échelles

Transformations admissibles	Types d'échelle	Exemple
ϕ o f(x) = f(x) (Identité)	Absolu	Le dénombrement
ϕ o f(x) = α f(x), $\alpha > 0$ Transformation similitude	Ratio	La masse
ϕ o f(x) = α f(x) + β, $\alpha > 0$ Transformation affine positive	Intervalle	La température (Fahrenheit, centigrades)
f(x) $\geq$ f(y) $\Leftrightarrow$ ϕ o f(x) $\geq$ ϕ o f(y) Transformation monotone positive	Ordinale	Une préférence
Toute fonction ϕ	Nominale	Une codification

3. La notion de signifiance

3.1. Définition

La notion de "*signifiance*" permet de comprendre ce qu'implique l'opération de mesurage et de connaître les traitements mathématiques autorisés pour un niveau de mesurage donné. C'est cette notion, issue de la théorie du mesurage, qui est principalement exploitée dans le cadre de cet article.

Commençons par introduire cette notion en prenant appui sur un exemple (Vansnick, 1991). Considérons un ensemble d'actions A = {a_1, a_2, a_3, a_4} et supposons que l'information fournie par le décideur soit la suivant e : il préfère a_1 à a_2, a_2 à a_3, a_1 à a_3, a_4 à a_3 et a_1 à a_4 ; enfin, il est indifférent entre a_2 et a_4. Cet ensemble d'informations peut être formalisé par une relation de préférence représentée figure 1 complétée par l'indifférence a_2 I a_4.

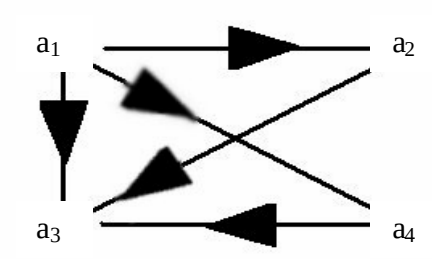

Figure 1 : Graphe des préférences

Cet ensemble d'informations peut être également représenté à l'aide de la fonction :

$$f : A \rightarrow R$$
$$a \rightarrow f(a) = |\{y \in A \,/\, aPy\}|$$

Ceci impliquerait les valeurs ci-après pour les éléments de X :

$$f(a_1) = 3,\ f(a_2) = 1,\ f(a_3) = 0,\ f(a_4) = 1.$$

Remarquons que nous avons, avec cette échelle de mesure, la relation suivante :

$$f(a_1) - f(a_2) = 2[f(a_4) - f(a_3)].$$

Considérons maintenant la proposition : "l'écart de préférence de a_1 par rapport à a_2 est deux fois plus grand que l'écart de préférence de a_4 par rapport à a_3". Nous dirons que cette proposition n'est pas signifiante. En effet, l'information obtenue du décideur peut aussi être représentée de la façon suivante :

$$H : A \rightarrow R \text{ tel que } h(a_1) = 10,\ h(a_2) = 8,\ h(a_3) = 0,\ h(a_4) = 8.$$

Nous observons que cette nouvelle fonction est encore une rna mais qu'elle invalide la relation précédente, c'est-à-dire que :

$$h(a_1) - h(a_2) = 2 \neq 2[h(a_4) - h(a_3)] = 2 \text{ x } 8.$$

Pour cette raison, nous pouvons dire que la proposition précédente n'est pas signifiante ; elle dépend de l'échelle de mesure choisie. La véracité de la proposition a changé car elle repose sur un calcul qui fait intervenir des caractéristiques des nombres que sont les échelons de l'échelle alors que ces derniers ont été définis sans lien avec une propriété concrète que le calcul présuppose pour donner un sens à cette caractéristique. En général, nous dirons **qu'une proposition fondée sur un calcul utilisant les échelons d'une échelle est signifiante si sa véracité ou sa fausseté demeure inchangée lorsqu'on remplace une échelle par une autre représentant toutes les deux la même information**. Cette définition, qui est due à Roberts (1979, p. 59), est étroitement liée au concept de l'invariance en géométrie. En fait, la notion de signifiance constitue une forme particulière de l'invariance dans la théorie du mesurage selon Falmagne et Narens (1983). D'autres définitions, légèrement différentes, ont été proposées dans la littérature. En effet, Suppes et Zinnes (1963) définissent la notion de signifiance de la manière suivante : "*A numerical statement is meaninful if and only if its truth (or falsity) is constant under admissible scale transformations of any of its numerical assignements, that is, any of its numerical functions expressing the results of*

measurement". Cette définition est plus précise que la précédente en ce sens qu'elle utilise le concept de transformation admissible des échelles ; par contre, elle est moins générale que la définition due à Roberts. En effet, Roberts et Franke (1976) ont montré que cette définition n'était valable que lorsqu'il s'agissait d'échelles régulières.

Dans cette étude, nous adoptons la définition de Suppes et Zinnes, les différentes échelles utilisées en analyse multicritère étant en général des échelles régulières. Pour terminer, il importe de souligner que **le concept de signifiance n'est pas synonyme de vérité.**

3.2. Illustration du concept de signifiance à partir de la somme pondérée

La somme pondérée constitue la procédure d'agrégation multicritère (PAMC) qui est, sans aucun doute, la plus familière et la plus employée. Elle est présente de façon implicite dans plusieurs méthodes d'évaluation globale. Elle est utilisée dans les conditions suivantes : chaque action potentielle a $\in$ A (ensemble des actions potentielles) est évaluée selon un ensemble d'attributs, de points de vue, de critères i, i =1, …, n, à partir d'échelles de mesure g[6]. On supposera dans tout ce qui suit que la préférence croît avec la performance $g_i(a)$. En tenant compte du fait que tous ces critères n'ont pas la même importance, on cherche à évaluer globalement chaque action a et à conclure si l'action a est ou non préférée à l'action b.

Avec la somme pondérée, on conclura que l'action a est préférée à b si on a :

$$\sum_{i=1}^{n} p_i g_i(a) > \sum_{i=1}^{n} p_i g_i(b) \qquad (3.1)$$

où $p_1, p_2, \ldots, p_n \in R^+$ représentent les poids des différents points de vue.

Or, il est important d'examiner dans quelles conditions l'action a reste préférée à l'action b lorsqu'on substitue aux échelles g des échelles "équivalentes" f. Il s'agit autrement dit d'étudier les conditions dans lesquelles la proposition (3.1) est signifiante. L'étude de cette conclusion vis-à-vis de la signifiance s'avère d'un grand intérêt car elle constitue une justification pour la prise de décision finale. En effet, tel que Roberts (1979) l'écrit : "*If a decision based on a numerical scale can be reversed by simply modifying the numbers, and*

[6] Attirons l'attention sur cette façon de désigner une échelle qui consiste à ne nommer que l'application : elle est assez elliptique puisqu'elle laisse sous-entendus aussi bien l'ensemble des échelons susceptibles d'être utilisés pour guider l'information que les propriétés concrètes que ce codage est censé représenter. Cette façon de s'exprimer est néanmoins commode et nous l'utiliserons chaque fois qu'elle ne prête pas à ambiguïté. Dans certains cas, il peut être préférable de désigner l'échelle par l'ensemble des échelons qui lui sont associés en mettant en évidence les propriétés que le codage lié à l'application (qui, s'il n'y a pas ambiguïté, peut rester sous-entendu) est censé représenter (cf. Roy, 1999).

if both sets of numbers legitimately represent the decision problem, then we have no good reason to make one decision rather than the other".

Nous allons étudier la signifiance de la proposition (3.1) tout d'abord pour des échelles toutes de plus haut niveau (échelles absolues) puis en descendant dans la hiérarchie du tableau 1. En effet, d'une façon générale, nous pouvons affirmer que si une procédure n'est pas signifiante pour un niveau de mesure donné, elle ne le sera pas pour un niveau de mesure "plus bas".

Il est évident que la proposition (3.1) est signifiante si les g_i sont des échelles absolues. En fait, pour ce type d'échelle, le problème de la signifiance ne se pose pas, car celle-ci n'admet aucune transformation. Plus généralement, toutes les propositions ne faisant intervenir que des échelles absolues sont signifiantes.

Considérons maintenant le cas où toutes les échelles sont des échelles ratio. Nous savons que les transformations admissibles pour une échelle ratio sont de la forme $\phi \circ g(a) = \alpha \cdot g(a)$, $\alpha > 0$. Ainsi, une échelle équivalente à g_i s'écrira sous la forme $h_i(a) = \alpha_i * g_i(a)$, $\alpha_i > 0$. Pour que la proposition (3.1) soit signifiante, il faudrait que la condition suivante soit satisfaite quelles que soient les valeurs des $\alpha_i > 0$:

$$\left[\sum_{i=1}^{n} p_i g_i(a) > \sum_{i=1}^{n} p_i g_i(b)\right] \Leftrightarrow \left[\sum_{i=1}^{n} p_i a_i g_i(a) > \sum_{i=1}^{n} p_i a_i g_i(b)\right],$$

ce qui n'est évidemment pas possible.

Nous pourrions ainsi être amenés à conclure que la procédure "somme pondérée" n'est pas signifiante pour des échelles de mesure de niveau ratio et qu'elle ne l'est pas non plus pour des échelles intervalles et encore moins pour des échelles de niveau ordinal.

Ce résultat fait ressortir un "piège" souvent insoupçonné par les utilisateurs de la somme pondérée, à savoir que les coefficients de pondération p_i dépendent des échelles de mesure g_i associées aux critères (Roy et Mousseau, 1996). Il n'est donc pas possible dans ce cas de se prononcer sur les valeurs des p_i sans prendre en compte les codifications numériques de ces échelles puisque ces p_i, qui sont des constantes d'échelle, jouent un rôle de compromis entre les différents critères.

Par exemple, si l'on est en présence de deux critères l'un de coût et l'autre de délai, les valeurs des p_i, i = 1, 2, ne seront pas les mêmes selon que les coûts sont exprimés en $, en

milliers de $ ou en millions de $ et que les délais sont exprimés en jours, en semaines ou en mois.

Ainsi, avec des échelles de niveau ratio, la somme pondérée sera signifiante si la transformation $h_i(a) = \alpha_i \cdot g_i(a)$ est accompagnée d'une transformation concomitante des p_i, c'est-à-dire transformant p_i en p_i/α_i. **La "somme pondéré" est donc signifiante sous réserve de transformation des coefficients de pondération de manière appropriée**. Ce résultat reste vrai pour des échelles de niveau intervalle. Dans ce cas, les transformations admissibles sont de la forme :

$$h_i(a) = a_i g_i(a) + b_i, a_i > 0.$$

La proposition (3.1) reste bien signifiante pour de telles échelles puisque :

$$\sum_{i=1}^{n} \frac{P_i}{a_i}[a_i \, g_i(a) + b_i] > \sum_{i=1}^{n} \frac{P_i}{a_i}[a_i \, g_i(b) + b_i] \Leftrightarrow$$

$$\sum_{i=1}^{n} \frac{P_i}{a_i} g_i(a) + \frac{P_i b_i}{a_i} > \sum_{i=1}^{n} \frac{P_i}{a_i} g_i(b) + \frac{P_i b_i}{a_i} \Leftrightarrow$$

$$\sum_{i=1}^{n} P_i g_i(a) > \sum_{i=1}^{n} P_i g_i(b)$$

Souvent, les performances telles qu'elles s'expriment avec l'échelle originelle (concrète) donnent lieu à une normalisation avant agrégation. Cette normalisation vise généralement à s'affranchir des unités propres à chacune de ces échelles originelles. Elle peut par exemple permettre de travailler avec des performances toutes comprises dans l'intervalle 0-1 ou 0-100. Les techniques de normalisation sont nombreuses (cf. Maystre *et al.*, 1994). Avec une agrégation par somme pondérée, la détermination des coefficients de pondération p_i qui s'appliqueront aux performances normalisées doit être conduite de façon à garantir la signifiance de la proposition (3.1). Nous reviendrons sur ce sujet au 4.1.

Considérons enfin le cas où l'une des échelles au moins est seulement ordinale. Il est clair que la proposition (3.1) n'est plus signifiante. On peut cependant encore utiliser l'agrégation par somme pondérée pour formaliser des propositions signifiantes. C'est notamment le cas lorsque celles-ci portent non plus sur des performances mais sur les relations que ces performances expriment relativement à chaque critère sur l'ensemble des actions. Posons par exemple :

$$\upsilon_i(a, b) = 1 \text{ si } a \, P_i \, b \text{ et } \upsilon_i(a, b) = 0 \text{ si non } a \, P_i \, b.$$

Les échelles υ_i qui sont ainsi définies pour tout élément du produit cartésien $A \times A$ sont des échelles absolues. Il s'ensuit qu'une proposition caractérisée par la formule suivante (afin de comparer par exemple a à b) :

$$\sum_{i=1}^{n} p_i n_i(a,b) > \sum_{i=1}^{n} p_i n_i(b,a)$$

est signifiante.

4. Analyse de quelques PAMC

Dans cette section, nous analysons un ensemble de PAMC du point de vue de la signifiance. Selon Roy et Bouyssou (1993) "*une PAMC est une règle, un procédé permettant d'établir, sur la base du tableau des performances et d'informations inter-critères (des coefficients d'importance, des seuils de veto, des taux de substitution, …), un ou plusieurs systèmes relationnels de préférences (SRP) sur l'ensemble A des actions (et, éventuellement, sur l'ensemble des mutations (A ´ A))*".

Le SRP sur A dont il est question dans cette définition peut prendre des formes diverses : (I, P), (I, P, R), (S, R), … où R désigne la relation d'incomparabilité. Le tableau des performances contient l'ensemble des évaluations des actions selon chacun des critères/attributs plus, éventuellement, des seuils de discrimination.

Les procédures analysées dans cette section appartiennent aux deux approches opérationnelles qualifiées d'approche du critère unique de synthèse et d'approche du surclassement de synthèse.

4.1. PAMC de type critère unique de synthèse

4.1.1. Généralités

La PAMC somme pondérée constitue un cas particulier des PAMC étudiées dans cette section. Comme leur nom l'indique ces PAMC visent à établir un SRP sur l'ensemble des actions en bâtissant un critère unique de synthèse

$$g(a) = G(g_1(a),..., g_i(a),..., g_n(a))$$

où g(a) est la performance de l'action a selon le critère ou l'attribut i.
Le SRP généralement associé au critère g est de type (I, P) avec :

$$a\,I\,b \Leftrightarrow g(a) = g(b),\ a\,P\,b \Leftrightarrow g(a) > g(b). \qquad (4.1)$$

Toute incomparabilité est donc éliminée et les relations d'indifférence I et de préférence P sont transitives. C'est donc essentiellement la signifiance de ces propositions d'indifférence et de préférence que nous introduirons ci-après. On peut ne s'intéresser qu'à celles concernant la préférence puisque l'indifférence a lieu si et seulement si la préférence n'a pas lieu.

On analysera successivement dans cette section les critères de synthèse associés aux procédures MAUT, SMART et TOPSIS. Nous laisserons notamment de côté la procédure AHT (cf. Saaty, 1977, 1984) qui a donné lieu à de multiples controverses. La position de Barzilai (1998) mérite toutefois d'être citée : "*A corollary is that procedures which measure preferences on a unit-scale in the absence of a universal or absolute zero (such as the Analytic Hierarchy Process) operate on an incorrect type of scale and consequently, produce incorrect results"* (voir aussi Bana e Costa et Vansnick, 2001).

4.1.2. MAUT (Keeney et Raiffa, 1976)

La forme analytique du critère de synthèse découle ici de la théorie de l'utilité multiattribut (MAUT). Dans cette théorie (qui repose sur un corps d'axiomes et d'hypothèses extrêmement précis), chaque critère g se présente comme une espérance d'utilité faisant intervenir une fonction d'utilité u_i relative à un attribut (ou un groupe d'attributs dans certains cas) x_i. Pour chaque action a, cet attribut est caractérisé par sa distribution de probabilités, laquelle peut, le cas échéant, refléter la certitude[7].

Dans cette théorie, le système relationnel de préférences (I, P) qu'il s'agit de modéliser est supposé préexister dans l'esprit du décideur. Lorsqu'il est conforme aux axiomes et hypothèses de la théorie, celle-ci prouve qu'il peut être fidèlement représenté conformément aux relations (4.1) avec un critère de synthèse g (correspondant à une espérance d'utilité). Sous réserve de supposer vérifiées des hypothèses d'indépendance au sens des utilités et au sens des préférences, ce critère de synthèse prend alors nécessairement l'une ou l'autre des deux formes suivantes (la première étant une forme limite de la seconde) :

Forme additive : $g(a) = \sum_{i=1}^{n} k_i \, g_i(a)$ avec $k_i > 0, \ \sum_{i=1}^{n} k_i = 1$

Forme multiplicative : $g(a) = \dfrac{\prod_{i=1}^{n} [1 + K \, k_i \, g_i(a)] - 1}{K}$ (4.2)

[7] La version non probabiliste de MAUT (MAVT) nous paraît une démarche rigoureuse pour bâtir un critère de synthèse de type somme pondérée (cf. § 3.2).

avec $k_i > 0, \sum_{i=1}^{n} k_i \neq 1, \quad K > -1, \quad K \neq 0,$

K étant l'unique racine réelle non nulle et supérieure à – 1 de l'équation :

$$\prod_{i=1}^{n} [1 + K\, k_i] = 1 + K.$$

Avec chacune de ces deux formes, il faut souligner que l'origine du critère g est définie par l'action (fictive) a_* qui vérifie $g_i(a_*) = 0$ pour i = 1, ..., n et l'unité de ce critère est définie par l'action (fictive) a^* qui vérifie $g_i(a^*) = 1$ pour i = 1, ..., n.

Dès l'instant où l'on se place dans le cadre des hypothèses d'indépendance rappelées ci-dessus avec des attributs x_i dont le mesurage est de type ordinal (échelle non nécessairement de type intervalle), la théorie de l'utilité multiattribut établit que les fonctions d'utilité u_i existent et sont définies de façon unique à une transformation affine près. Il s'ensuit que le modélisateur est libre de choisir comme il l'entend l'origine et l'unité de chacune d'elles. Ces choix consistent habituellement à poser $u_i = 0$ à l'une des extrémités de l'échelle de mesure de x_i et $u_i = 1$ à l'autre extrémité. L'échelle de mesure du critère g_i est donc de type intervalle. Elle devient de type absolu dès l'instant où les choix de cette origine et de cette unité ont été effectués (peu importe comment). Ces choix étant faits pour chaque critère, la théorie affirme qu'il existe une valeur parfaitement définie pour chacun des coefficients k_i qui garantit :

$$g(a) > g(b) \Leftrightarrow a\ P\ b, \forall\ a, b \in A.$$

Ainsi, la proposition (4.1) est signifiante quelle que soit la façon dont ont été fixées les origines et les unités des fonctions u_i. Cela suppose bien évidemment que les valeurs des coefficients k_i aient été attribuées en conformité avec la théorie. Or, ces valeurs dépendent précisément de la façon dont ces origines et unités ont été fixées (c'est pourquoi on les appelle des constantes d'échelle). Considérons en effet l'action fictive b_i caractérisée par :

$$g_j(b_i) = 0, \forall\ j \neq i \text{ et } g_i(b_i) = 1.$$

Le lecteur vérifiera sans peine que $g(b_i) = k_i$ aussi bien avec la forme additive qu'avec la forme multiplicative. De plus, le choix entre la forme additive et la forme multiplicative est déterminé (cf. formules (4.2)) par la valeur que prend la somme des k_i. Il est donc essentiel, pour attribuer une valeur à ces constantes d'échelle, de respecter un protocole d'interrogation du détenteur du système relationnel de préférences (I, P) qui prenne en compte, conformément

à la théorie, la façon dont les critères g_i ont été normés (choix de l'origine et de l'unité de chacune des fonctions d'utilité u_i).

Les protocoles ainsi préconisés dans la littérature font intervenir des actions fictives comme par exemple a_*, a^*, b_i (définies ci-dessus) qui sortent généralement du champ de perception normal de la personne interrogée. On peut voir là une limite à la signifiance. C'est en effet une source de fragilité des réponses. De surcroît, on peut penser que la sensibilité des réponses aux conditions de normalisation est d'autant plus accusée que le système de préférences de la personne interrogée est en fait moins complètement préexistant. On peut supposer que le questionnement contribue, dans bien des cas, à construire pour une part ce système de préférence (cf. Roy, 1987).

En conclusion, la théorie de l'utilité multiattribut fournit ici un cadre de référence dans lequel les propositions validées à partir des formules (4.2) sont signifiantes. Toutefois, en pratique, on est souvent amené à élaborer le critère de synthèse sans pouvoir s'assurer que les hypothèses de cette théorie sont satisfaites et surtout en étant contraint de s'écarter du protocole rigoureux d'interrogation du décideur que cette théorie préconise (voir McCord et de Neufville, 1982).

4.1.3. SMART (Edwards, 1977 ; Von Winterfield et Edwards, 1986)

Le critère de synthèse g est ici du type somme pondérée :

$$g(a) = \sum_{i=1}^{n} k_i\, g_i(a) \text{ avec } \sum_{i=1}^{n} k_i = 1 \text{ et } 0 \leq g_i(a) \leq 100. \qquad (4.3)$$

Comme le montre cette formule, les performances doivent ici être normalisées (cf. § 3.2). Cette normalisation découle dans SMART de la façon dont sont définies les échelles.

- Si l'appréciation de la performance selon le critère i est de nature qualitative (avis d'experts par exemple), on doit coder 0 la plus mauvaise performance plausible et 100 la meilleure performance plausible, les appréciations intermédiaires devant être codées de façon à définir une échelle intervalle.

- Si l'appréciation de la performance selon le critère i est de nature numérique (ce qui ne signifie pas toujours objective et encore moins quantitative), on doit poser:

$$g_i(a) = \frac{h_i(a) - h_{i*}}{h_i^* - h_{i*}} \times 100 \qquad (4.4)$$

où $h_i(a)$ correspond à l'appréciation numérique originelle de l'action a et h_{i*}, h_i^* correspondent respectivement aux appréciations plausibles la plus mauvaise et la meilleure.

Il convient ici d'attirer l'attention sur un point important. Ce type de normalisation ne garantit nullement que les performances soient mesurées sur des échelles absolues au sens défini en section 2. Pour qu'il en soit ainsi, il est nécessaire (et d'ailleurs suffisant) que les appréciations originelles (avant normalisation) soient mesurées sur des échelles intervalle. Il y a là une condition **nécessaire** (mais non suffisante comme nous allons le voir) pour que, avec SMART, la proposition (4.1) soit signifiante. Dans bien des cas, rien ne prouve que cette condition soit satisfaite, même lorsque les appréciations initiales sont de nature numérique. En effet, l'échelle peut, dans ce cas aussi, être de type seulement ordinal (cf. § 2.2).

Le protocole préconisé par SMART pour attribuer des valeurs numériques aux coefficients k_i comporte deux étapes :

- "ranking" : rangement des critères par ordre d'importance croissante ou décroissante ;
- "rating" : mesurage de cette importance sur une échelle qui doit être de type ratio.

Pour que la proposition (4.1) soit signifiante, il est nécessaire que la validation de a P b ne dépende pas des conditions de normalisation. Il faut autrement dit qu'une modification éventuelle de la définition des appréciations de référence la plus mauvaise et la meilleure plausibles ne puisse pas affecter cette validation. Les coefficients k_i sont donc ici encore des constantes d'échelle. Si, pour une raison quelconque, on est amené à doubler la différence $h_i^* - h_{i*}$, alors le coefficient k_i doit être multiplié par deux (cf. § 3.2). Pour attribuer une valeur à ces constantes d'échelle qui garantisse la signifiance de la proposition (4.1) avec SMART, il est donc également **nécessaire** de faire intervenir explicitement, aussi bien au niveau du "ranking" qu'au niveau du "rating", ces appréciations de référence. C'est là une seconde condition qui est loin d'être toujours clairement remplie.

En conclusion, l'emploi des formules (4.3) et (4.4) nécessite, comme le mettent en évidence les deux conditions nécessaires ci-dessus, de prendre de grandes précautions dans leur mise en œuvre pour valider des propositions signifiantes. La méthode MACBETH (cf. Bana e Costa et Vansnick, 1994, 1997) fournit une démarche accompagnée des outils nécessaires pour utiliser l'agrégation par somme pondérée dans l'esprit ci-dessus de telle sorte que ces deux conditions soient satisfaites.

4.1.4. TOPSIS (Hwang et Yoon, 1981)

La définition du critère de synthèse g(a) repose, dans cette méthode, sur une "distance euclidienne" relative devant permettre de positionner chaque action a ∈ A par rapport à deux actions fictives dites *idéales* et *anti-idéales*. Si l'on note respectivement ces distances $d^*(a)$ et $d_*(a)$, le critère g est défini par :

$$g(a) = \frac{d_*(a)}{d_*(a) + d^*(a)}. \qquad (4.5)$$

Les actions idéales a^* et anti-idéales a_* sont caractérisées par les appréciations extrêmes que prennent les performances des actions de A avec les critères g_i :

$$d^*(a) = \left(\sum_{i=1}^{n} \frac{k_i^2}{g_i^2(A)} \left[g_i(a^*) - g_i(a) \right]^2 \right)^{\frac{1}{2}}. \qquad (4.6)$$

$$d_*(a) = \left(\sum_{i=1}^{n} \frac{k_i^2}{g_i^2(A)} \left[g_i(a) - g_i(a_*) \right]^2 \right)^{\frac{1}{2}}. \qquad (4.7)$$

avec

$$g_i^2(A) = \sum_{a \in A} g_i^2(a) \qquad (4.8)$$

où les coefficients k_i sont destinés à différencier l'importance du rôle que l'on souhaite faire jouer aux différents critères dans la PAMC.

Il convient ici d'attirer l'attention sur un point important. La formule (4.8) fait intervenir toutes les performances des actions selon le critère g_i. Il s'ensuit qu'une variation quelconque de la performance d'une action a donnée a_0 selon ce critère modifie la valeur de $g_i(A)$ et, par suite, les valeurs de $d^*(a)$ et $d_*(a)$ **pour toutes les actions a Î A**. Cet impact sur les distances pourrait être annulé par une variation concomitante (dans les mêmes proportions) des coefficients k_i. Autrement dit, faire varier la performance d'une action quelconque selon un critère équivaut à modifier l'importance donnée à ce critère. C'est là un piège dû à la façon dont la normalisation est conçue. En outre, la variation de $g_i(a_0)$ peut modifier la définition de a^* ou celle de a_* ; il peut en aller de même avec l'ajout ou le retrait d'une action. L'impact sur les distances $d^*(a)$ et $d_*(a)$ qui découle de ces transformations n'est alors plus équivalent à une modification des coefficients k_i.

Ces considérations suffisent à montrer que la proposition (4.1) n'est signifiante avec TOPSIS que si les coefficients k_i ont été raisonnés en fonction des valeurs de $g_i(A)$ et si les actions

idéales et anti-idéales apparaissent comme des références qui ne prêtent pas à discussion. Pour cette raison, les valeurs maximum et minimum plausibles introduites dans SMART pourraient paraître mieux adaptées que celles définies par les formules (4.6) et (4.7).

Si l'on fait abstraction des difficultés précédentes, il est facile de vérifier que la proposition (4.1) est signifiante dans TOPSIS avec des performances mesurées sur des échelles de niveau ratio. En effet, un changement d'unité (transformation $\alpha_i\ g_i$) laisse invariantes les distances $d^*(a)$ et $d_*(a)$ (cf. formules (4.6) et (4.7)). En revanche, un changement d'origine (transformation $g_i + \beta_i$) n'affecte que le terme $g_i(A)$ dans la formule (4.6). Il équivaut ici encore à une modification des coefficients k_i. La proposition (4.1) n'est donc pas signifiante avec des performances mesurées sur des échelles de niveau intervalle.

En conclusion, pour que l'emploi des formules (4.5)-(4.8) permette de valider des propositions signifiantes, il est en premier lieu nécessaire que les performances soient mesurées sur des échelles de niveau ratio. Il faut, en second lieu, avoir une compréhension très fine de la signification complexe que ces formules confèrent aux coefficients k_i ainsi qu'aux caractéristiques des actions idéales et anti-idéales pour pouvoir attribuer à ces paramètres des valeurs convenables et pour les interpréter correctement. Enfin, le fait que toute modification des performances d'une action doive s'accompagner d'un adaptation des coefficients k_i rend l'analyse de sensibilité et, *a fortiori*, l'analyse de robustesse quasiment impraticables.

4.2. PAMC de type surclassement de synthèse

4.2.1. Généralités

Nous analyserons ci-après trois de ces PAMC : celles sur lesquelles prennent appui les méthodes ELECTRE I, IS et III (voir Roy et Bouyssou, 1993 ; Roy, 1978) et PROMÉTHÉE I et II (voir Brans et Vincke, 1985 ; Vincke, 1989). Rappelons que, dans chacune de ces méthodes, la PAMC est complétée par une procédure d'exploitation dont l'objet est de tirer parti, pour l'aide à la décision, du SRP construit par la PAMC. Ce SRP tolère la présence d'incomparabilités qui sont mises en évidence :

- soit directement à l'aide d'une relation R d'incomparabilité comme c'est le cas dans les méthodes portant le nom ELECTRE ;
- soit indirectement à l'aide d'une relation ~ de non préférence (servant à modéliser l'indifférence et l'incomparabilité sans les différencier) comme c'est le cas dans les méthodes portant le nom PROMÉTHÉE.

Dans les méthodes PROMÉTHÉE, le SRP est de deux types (P, ~) où la relation (anti-symétrique) de préférence P n'est pas nécessairement transitive. Dans les méthodes ELECTRE, le SRP est du type (S, R) où la relation S (dite de surclassement) a la signification suivante :

> $\forall$ a, b $\in$ A : a S b $\Leftrightarrow$ a est globalement (compte tenu de tous les critères) au moins aussi bonne que b.

S n'est pas nécessairement transitive. Elle n'est pas non plus anti-symétrique : a S b et b S a modélisent les cas d'indifférence.

L'analyse de la signifiance dans chacune de ces PAMC porte donc, selon le cas, sur la façon dont est validée l'assertion "a surclasse b" ou "a est préférée à b" (l'incomparabilité ou la non préférence étant validée par complémentarité). Pour construire cette analyse et présenter, dans chaque cas, la PAMC concernée, nous nous conformerons aux notations de Roy et Bouyssou (1993, chapitre 5).

IV.2.2. ELECTRE I

La proposition "a surclasse b" (a S b) est ici validée si et seulement si deux conditions, dites l'une de **concordance**, l'autre de **non veto**, sont satisfaites. La première fait intervenir la coalition C(a, b) des critères en accord avec la proposition :

$$C(a, b) = \{i \,/\, g(a) \geq g_i(b)\}. \qquad (4.9)$$

On supposera que chacun des n critères est doté d'un "pouvoir de vote" caractérisé par le nombre k_i (> 0) pour le critère g. Posons :

$$\sum_{i=1}^{n} k_i = K \text{ (nombre de voix attribuées)}$$

Avec ces notations, la condition de concordance s'écrit :

$$c(a, b) = \frac{1}{K} \sum_{i \in C(a, b)} k_i \geq s, \quad \frac{1}{2} \leq s \leq 1. \qquad (4.10)$$

s est appelé le niveau de concordance : il s'interprète comme un niveau minimum de majorité requis pour valider la proposition. Faisons observer que $s > 1 - \min_i \frac{k_i}{K}$ implique l'unanimité.

Soit g_i un critère tel que $i \notin C(a, b)$. Un tel critère est en désaccord avec la proposition considérée puisque $g_i(b) > g_i(a)$. La méthode prévoit que, si le désaccord est "important", le critère g_i peut mettre son veto à la validation de la proposition a S b. Pour cela, il convient de caractériser l'ensemble des couples de performances (y, x) révélateurs d'une situation de désaccord jugés suffisamment critiques pour être regardés comme **incompatibles** avec le surclassement de b par a lorsque $g_i(b) = y$ et $g_i(a) = x$. Dans les méthodes ELECTRE, cet ensemble de couples est défini par une fonction $v_i(x)$ (> 0) dite seuil de veto. Les couples de performances révélateurs d'une situation critique sont caractérisés par ceux qui vérifient :

$$y - x > v_i(x).$$

Il découle de ces définitions que le critère g_i met son veto au surclassement a S b (du fait d'une situation critique incompatible avec la validation de cette proposition) si et seulement si :

$$g_i(b) - g_i(a) > v_i[g_i(a)]. \qquad (4.11)$$

On démontre que, pour des raisons élémentaires de cohérence, le seuil de veto $v_i(x)$ doit être tel que $x + v_i(x)$ soit une fonction monotone non décroissante de x.

La seconde des deux conditions nécessaires mentionnées plus haut pour valider la proposition a S b (dite condition de non veto) est satisfaite si et seulement si aucun des critères ne vérifie la condition (4.11).

Les échelles sur lesquelles sont mesurées les performances $g_i(a)$ sont au moins de niveau ordinal. Examinons le cas où elles ne sont pas d'un niveau supérieur. La proposition a S b est, dans ces conditions, signifiante si et seulement si sa validation demeure invariante lorsque l'un quelconque des critères, disons g_i, est remplacé par une autre codification $h_i = \phi(g_i)$ et cela quelle que soit la fonction monotone croissante ϕ. Montrons qu'il en est bien ainsi.

La substitution des h_i à g_i n'affecte en rien le fait que i appartienne ou n'appartienne pas à C(a, b) (formule (4.9)). Cette substitution ne modifie pas la valeur de c(a, b) (formule (4.10)). La condition de concordance est donc bien invariante. Il en sera de même avec la condition de non veto sous réserve de conserver la même définition des couples de situations critiques qui caractérisent les possibilités de veto. Si y et x constituent un couple de performances qui reflètent une telle situation avec le mesurage g_i, il va en être de même avec les performances

$\phi(y)$ et $\phi(x)$ obtenues avec le nouveau mesurage h_i. Il est donc nécessaire d'adapter la fonction seuil car celle-ci dépend du mesurage choisi. Ne pas faire cette adaptation équivaudrait à modifier les règles de validation du surclassement. Le lecteur vérifiera sans peine que la nouvelle fonction seuil $v_i^f(x)$ est définie sans ambiguïté par :

$$v_i^f[f(x)] = f[x + v_i(x)] - f(x). \qquad (4.12)$$

Il découle par exemple de cette formule que si $\phi(x)$ est la transformation affine $a x + ß$, la condition (4.11) doit s'écrire avec le nouveau mesurage :

$$h_i(b) - h_i(a) > \alpha\ v_i(a). \qquad (4.13)$$

En conclusion, la PAMC de ELECTRE I est signifiante avec des performances mesurées sur des échelles de niveau seulement ordinal et, a fortiori, de niveau supérieur. Il faut toutefois signaler que, avec des échelles de niveau supérieur, ces PAMC n'exploitent pas à fond l'information contenue dans les performances ainsi mesurées. En revanche, la façon de coder numériquement des appréciations de nature qualitative est sans effet sur la validation ou la non validation d'un surclassement.

4.2.3. ELECTRE IS et ELECTRE III

Ces deux PAMC font intervenir une famille de critères qui peut comporter des pseudo-critères. En premier lieu, nous rappellerons (cf. Roy, 1985) les principaux traits caractéristiques de ce concept. Ces PAMC prennent appui sur un même indicateur de concordance qui étend, au cas des pseudo-critères, la notion de majorité concordante introduite à propos d'ELECTRE I mais elles diffèrent dans leur façon de faire intervenir d'éventuels veto. Nous examinerons donc, en second lieu, le cas où aucun des critères n'est autorisé à mettre un veto. Nous étudierons enfin le cas général pour l'une puis pour l'autre de ces deux PAMC.

a) A propos du concept de pseudo-critère

Considérons un critère g conçu pour comparer des actions de A selon un point de vue spécifié. Il devient un pseudo-critère dès l'instant où l'on considère que, étant donné deux actions a et b ayant des performances $g(a) = x$ et $g(b) = y$ ($y \geq x$) :

- il peut y avoir indifférence entre a et b selon ce critère alors que $y > x$: l'ensemble des couples d'échelons (x, y) pour lesquels il en est ainsi doit alors être caractérisé comme suit :

$$x \leq y \leq x + q(x) ;$$

q(x) est appelé **seuil d'indifférence** ;

– il n'y a préférence **stricte** de b vis-à-vis de a que si l'écart entre les performances y et x est jugé significatif : l'ensemble des couples d'échelons pour lesquels il en est ainsi doit alors être caractérisé comme suit :

$$x + q(x) \leq x + p(x) < y \text{ ;}$$

p(x) est appelé **seuil de préférence**.

Soulignons que les deux fonctions-seuils sont définies sur l'ensemble des échelons utilisé pour coder l'information. Elles ont pour raison d'être le fait qu'une différence $y - x \neq 0$ peut ne pas constituer un argument probant pour affirmer que, selon le point de vue considéré, b est incontestablement préféré à a. Cela peut provenir pour une part de la nature même des échelons jugés trop proches pour que leur différence puisse être regardée comme apte à révéler une préférence et, pour une autre part (généralement beaucoup plus déterminante), de la nature de la représentation numérique g. Cette dernière, en effet, ne conduit pas, en général, à un mesurage x de l'action a qui soit indiscutable (et de même avec y pour l'action b). Cela tient au fait que ce que l'on cherche à mesurer est très souvent imparfaitement déterminé. Cette mauvaise détermination peut avoir des origines variées : l'information de base qui a été codée peut être purement verbale, elle peut provenir d'une enquête, elle peut faire appel à une formule analytique discutable, elle peut dépendre de paramètres dont la valeur reste subjective, ... Compte tenu de cette mauvaise détermination, l'action a aurait pu tout aussi bien être mesurée avec l'échelon x' et b avec l'échelon y' tels que y' soit légèrement inférieur à x' alors que y était légèrement supérieur à x.

Avec deux seuils, le pseudo-critère offre la possibilité d'avoir (pour certaines valeurs de x) $q(x) < p(x)$. Ceci permet d'introduire, dans la modélisation, une forme d'ambiguïté : il s'agit de celle qui découle du caractère fréquemment non nettement tranché entre une indifférence franchement acceptée et une préférence solidement argumentée. Un couple (x, y) tel que :

$$q(x) < y - x \leq p(x)$$

correspond alors à une hésitation entre indifférence et préférence stricte. Il s'agit d'une préférence dite **faible** (au sens de faiblement établie). Elle conduit à exclure la préférence stricte de a sur b sans qu'il soit possible d'être plus précis.

Soulignons, pour terminer ces rappels, que, pour les mêmes raisons de cohérence que celles invoquées à propos du seuil de veto (cf. 4.2.2), les deux fonctions x + q(x) et x + p(x) doivent être monotones non décroissantes.

b) Absence de veto

En l'absence de veto, seule la valeur que prend un indicateur de concordance détermine :

- dans ELECTRE IS la validation ou la non validation de la proposition *a S b*;
- dans ELECTRE III le niveau de crédibilité de la proposition *a S b*.

Dans un cas comme dans l'autre, la PAMC est donc complètement définie par la valeur de cet indicateur que nous noterons encore *c(a, b)*. Pour le définir, il nous faut introduire non plus une seule mais deux coalitions (il s'agit de deux sous-familles disjointes de F) :

$$C(a\ S\ b) = \{j\ /\ j\ \hat{I}\ F,\ g_j(a) + q_j(g_j(a))\ ^3\ g_j(b)\},$$
$$C(b\ Q\ a) = \{j\ /\ j\ \hat{I}\ F,\ g_j(a) + q_j(g_j(a)) < g_j(b)\ £\ g_j\ (a) + p_j(g_j(a))\}. \qquad (4.14)$$

On notera ici encore k_j le pouvoir de vote attribué au critère g_j et $K = \sum_{j \in F} k_j$ le nombre total des voix ainsi attribuées. L'indice de concordance c(a, b) est défini par :

$$c(a, b) = \frac{1}{K} \sum_{j \in C(aSb)} k_j + \frac{1}{K} \sum_{j \in C(bQa)} j_j(a, b)\ k_j \qquad (4.15)$$

avec $j_j(a, b) = \dfrac{g_j(a) + p_j(g_j(a)) - g_j(b)}{p_j(g_j(a)) - q_j(g_j(a))}$. (4.16)

Faisons observer que, *" j Î C(b Q a) : $p_j(g_j(a)) > q_j(g_j(a))$ et 0 £ $j_j(a, b)$ £ 1*. Considérons tout d'abord le cas où $q_j(x) = p_j(x)$, ∀ j ∈ F et ∀ x (on parle dans ce cas de quasi-critère). On a alors *C(b Q a)* = ∅. Plaçons-nous au plus bas niveau des échelles de mesure, à savoir échelles seulement ordinales (cf. 2.2). Les deux PAMC étudiées seront signifiantes pour ce type d'échelles si et seulement si *c(a, b)* est invariant lorsqu'on substitue à l'un quelconque des critères g_i la codification critère $h_i = f_i(g_i)$ avec f_i fonction monotone strictement croissante quelconque. Dans les conditions considérées ici, il en est bien ainsi. En effet, l'appartenance ou la non appartenance de i à la coalition *C(a S b)* n'est pas affectée par une telle substitution. Cela découle du fait que l'ensemble des couples d'échelons compatibles avec l'indifférence doit bien évidemment rester le même : le changement d'échelle s'accompagne (comme dans le cas du seuil de veto) d'une adaptation du seuil d'indifférence. La nouvelle fonction-seuil $q_i^f(x)$ est liée à l'ancienne $q_i(x)$ par la formule (4.12) dans laquelle q remplace v.

Les conclusions formulées à la fin du 4.2.2 pour la PAMC d'ELECTRE I restent donc valables pour celles d'ELECTRE IS et ELECTRE III lorsque les pseudo-critères ne sont que des quasi-critères et en l'absence de veto.

Considérons maintenant le cas général où (toujours en l'absence de veto) certains des critères sont effectivement des pseudo-critères (et non pas simplement des quasi-critères). Pour les mêmes raisons que celles indiquées ci-dessus, les coalitions $C(a\ S\ b)$ et $C(b\ Q\ a)$ sont invariantes lorsque l'on substitue, au critère g_i, la codification $h_i = f_i(g_i)$ (sous réserve de bien adapter les deux fonctions-seuils). Ce changement d'échelle peut néanmoins avoir un impact sur la valeur de $c(a, b)$ si (et seulement dans ce cas) i Î $C(b\ Q\ a)$. Il est facile de démontrer que (en transposant, aux seuils d'indifférence et de préférence, le résultat mis en évidence par la formule (4.13) à propos du seuil de veto) la valeur de $f_j(a, b)$ (cf. formule (4.16)) est invariante si f_i est une transformation affine ; la valeur de $c(a, b)$ est alors elle aussi invariante.

En conclusion, **c(a, b) n'est affecté par un changement d'échelle que si celui-ci n'est pas de type affine et s'il concerne un pseudo-critère vis-à-vis duquel la comparaison des actions a et b donne lieu à ambiguïté. Il s'ensuit que l'atteinte à la signifiance qui en découle pour les deux PAMC considérées (en l'absence de veto) peut être regardée comme le reflet de cette ambiguïté**.

c) ELECTRE IS avec veto

Pour valider la proposition $a\ S\ b$, il faut maintenant faire intervenir, conjointement avec la condition de concordance $c(a, b)$ ³ s, le fait qu'aucun critère ne met son veto à la validation de cette proposition. Dans ELECTRE IS, le critère g_i met son veto à la proposition a S b si et seulement si :

Condition de veto : $g_i(b) - g_i(a) > v_i(g_i(a)) - q_i(g_i(b))\dfrac{1 - c(a, b) - k_i/K}{1 - s - k_i/K}$. (4.17)

Faisons remarquer que si $q_j(x) = p_j(x) = 0$, ∀ j et ∀ x, alors les conditions de concordance et de veto d'ELECTRE IS se ramènent à celles d'ELECTRE I. En outre, si $q_i(g_i(b)) = 0$, le terme qui, dans (4.17) vient renforcer l'effet de veto (et cela d'autant plus que $c(a, b)$ se rapproche du niveau minimum de majorité requis s) disparaît. Quoi qu'il en soit, ce terme (qui constitue la seule modification à la formule (4.11)) est invariant dans un changement d'échelle dès l'instant où $c(a, b)$ l'est. Autrement dit, si $c(a, b)$ n'est pas affecté par un changement d'échelle du critère g_i, alors la condition de veto liée à ce critère ne le sera pas non plus.

Les conclusions énoncées au b) ci-dessus relativement à la signifiance de la PAMC ELECTRE IS restent les mêmes lorsqu'il existe des possibilités de veto.

d) ELECTRE III avec veto

Afin de prendre en compte la possibilité donnée à un pseudo-critère g_j de mettre son veto au surclassement $a\ S\ b$, la PAMC d'ELECTRE III fait intervenir **un indice de discordance** $d_i(a,b)$. Cet indice vise à traduire la plus ou moins grande opposition à une validation de $a\ S\ b$ que révèle un écart positif $g_i(b) - g_i(a)$. Dès l'instant où cet écart est suffisamment grand pour qu'il y ait veto, l'indice vaut 1 (valeur maximum). En revanche, il est nul si, avec le critère considéré, on a non $b\ P_i\ a$ (cf. a) ci-dessus). Enfin :

$$\text{si } p_i(g_i(a)) \le g_i(b) - g_i(a) \le v_i(g_i(a))$$

$$\text{alors } d_i(a,b) = \frac{g_i(b) - g_i(a) - p_i(g_i(a))}{v_i(g_i(a)) - p_i(g_i(a))} \qquad (4.18)$$

Rappelons que (cf. b) ci-dessus) lorsque le droit de veto n'est pas reconnu à aucun critère, la crédibilité de l'assertion $a\ S\ b$ vaut $s_S(a,b) = c(a,b)$ (cf. formules (4.14), (4.15), (4.16)). Cette crédibilité est affaiblie par un critère g_i discordant auquel on a reconnu un droit de veto si et seulement si $d_i(a,b) > c(a,b)$. Cet affaiblissement annule la crédibilité dès l'instant où $d_i(a,b) = 1$. De façon plus précise :

$$s_S(a,b) = c(a,b) \times \prod_{i \in D_c(a,b)} \frac{1 - d_i(a,b)}{1 - c(a,b)}$$

$$\text{avec } D_c(a,b) = \{i / i \in F, d_i(a,b) \ge c(a,b)\} \qquad (4.19)$$

Examinons maintenant l'impact que peut avoir, sur $s_S(a,b)$. Le remplacement de certains critères g_i par $h_i = f_i(g_i)$ pour i $\in H$. Alors que, dans les PAMC examinées jusqu'à maintenant, on pouvait se contenter (en raison de la forme additive des formules) d'analyser séparément, critère par critère, l'impact de ce type de changement, il peut se produire, dans la formule (4.19), des phénomènes d'interaction qui obligent à raisonner sur un sous-ensemble H de critères non nécessairement réduits à un seul. Pour chacun des critères pris en compte, il convient bien évidemment d'adapter non seulement les seuils d'indifférence et de préférence mais aussi le seuil de veto. Il découle de cette adaptation que si $d_i(a,b) = 0$ ou 1 avec g_i, cette valeur restera la même avec h_i quelle que soit la transformation monotone f_i. Elle pourra toutefois être modifiée si $0 < d_i(a,b) < 1$.

Considérons tout d'abord le cas où toutes les transformations f_i sont de type affine. Comme nous l'avons vu au b) ci-dessus, $c(a,b)$ est alors invariant. $d_i(a,b)$ l'est aussi même pour des critères qui vérifient les inégalités (4.18) (ceci découle toujours du résultat mis en évidence par la formule (4.13) à propos du seuil de veto). Il s'ensuit que, quel que soit H, $D_c(a,b)$ est aussi invariant. Il en va donc de même pour $s_S(a,b)$.

Considérons maintenant le cas où une seule des transformations f_i n'est pas affine. Soit g_i le critère concerné. Trois cas exclusifs sont seuls possibles.

- $j \in C(a\ S\ b)$ (formule (1.14)) : $c(a,b)$ est invariant et il en est donc de même pour $s_S(a,b)$.
- $j \in C(b\ Q\ a)$ (formule (1.14)) : la variation de $c(a,b)$ est égale à celle de $f_i k_i / K$ (qui en général, sera négligeable et de toute façon bornée par k_i / K : la variation de $s_S(a,b)$ qui en découle pourra être amplifiée par celle des fractions $\dfrac{1-d_i(a,b)}{1-c(a,b)}$, $\forall \quad i \in D_c(a,b)$ (qui, si cet ensemble est non vide, vraie dans le même sens que $c(a,b)$ ainsi que par la possible modification de la composition de l'ensemble $D_c(a,b)$; rappelons que le cas considéré ici suppose qu'il existe une zone d'ambiguïté non nulle caractérisée par l'intervalle $\lfloor q_j(g_j(a)), p_j(g_j(a)) \rfloor$.
- $d_i(a,b) \neq 0$: si $d_i(a,b) \leq c(a,b)$ ou si $d_i(a,b) = 1$, le remplacement de g_i par h_i laissera $s_S(a,b)$ invariant (formule (4.19)) ; ce remplacement n'aura donc un impact que si, avec l'un au moins de ces deux mesurages, $c(a,b) < d_i(a,b) < 1$; lorsqu'il en est ainsi, il paraît légitime de considérer que cette variation ne fait que refléter les difficultés qu'il y a à appréhender la crédibilité de l'assertion $a\ S\ b$ compte tenu du caractère imparfait de l'information.

Il reste enfin à examiner le cas où il est justifié de prendre en considération plus d'une transformation f_i de forme non affine. Cela suppose qu'il existe plus d'un critère faisant intervenir une échelle de niveau inférieur au niveau intervalle. Les résultats mis en évidence ci-dessus font apparaître des variations possibles de $s_S(a,b)$ seulement si, pour tels critères $q_j(g_j(a)) \neq p_j(g_j(a))$ ou si le droit au veto leur est reconnu. Dans l'un et l'autre cas, les variations de crédibilité devraient rester en général faibles si l'on est en présence d'échelles

d'intervalle faible car ce type d'échelle restreint les domaines des transformations f_i admissibles. Quoi qu'il en soit, ces variations doivent être regardées comme le reflet de difficultés inhérentes à la qualité des données pour appréhender un indice de crédibilité du surclassement.

En conclusion, **dans ELECTRE III, l'indice de crédibilité de l'assertion $a\ S\ b$ n'est affecté par un changement d'échelle que si celui-ci n'est pas de type affine. Il s'ensuit que cette PAMC est signifiante lorsque toutes les échelles sont de niveau intervalle : lorsque certaines d'entre elles sont de niveau inférieur, la possible atteinte à la signifiance peut être regardée comme le reflet de difficultés inhérentes à la qualité des données pour modéliser les préférences.**

4.2.4. PROMÉTHÉE I et II

L'analyse de la signifiance dans ces méthodes se réduit (comme dans ELECTRE III) à celle d'un indice de crédibilité (cf. 4.2.1). Cet indice $s_P(a,b)$ concerne ici la crédibilité de l'assertion $a\ P\ b$. Il est défini comme suit :

$$s_P(a,b) = \frac{1}{K}\sum_{j\in F} k_i s_{P_j}(a,b) \text{ avec } k_j \geq 0, \sum_{j\in F} k_j = K \qquad (4.20)$$

où $s_{P_j}(a,b)$ traduit la crédibilité de l'assertion $a\ P_j\ b$ (considérée du seul point de vue du critère g_i). Ces indices de crédibilité partielle $s_{P_j}(a,b)$ s'annulent dès l'instant où $j \in C(b\ S\ a)$ (cf. formule (4.14)). Il s'ensuit que l'indice de crédibilité global $s_P(a,b)$ ne permet pas de prendre en compte d'éventuels effets de veto. Il reflète la plus ou moins grande concordance des critères de F avec l'assertion $a\ P\ b$.

La définition précise de l'indice partiel s_{P_j} dépend de la nature du critère g_i. Dans PROMÉTHÉE en effet, six types de critères sont proposés. Les quatre premiers reposent sur la prise en compte de seuils d'indifférence q_j et de préférence p_j, tous les deux constants, qui peuvent être soit nuls, soit non nuls. Pour chacun d'eux, l'indice de crédibilité est défini par la formule générale suivante :

$$s_{P_j}(a,b) = \begin{cases} 0 & si \quad g_j(a) - g_j(b) \geq q_j \\ g_j(a) - g_j(b) - q_j & si \quad q_j < g_j(a) - g_j(b) \leq p_j \\ 1 & si \quad g_j(a) - g_j(b) > p_j \end{cases} \qquad (4.21)$$

Examinons tout d'abord la signifiance de $s_P(a,b)$ relativement à un critère g_i relevant de l'un de ces quatre premiers types. L'analyse conduite au 4.2.3 b) pour étudier l'impact sur l'indice de concordance $c(a,b)$ du remplacement de g_i par $h_i = f(g_i)$ peut être conduite, dans les mêmes termes, pour étudier l'impact d'un tel remplacement sur s_{P_j} et sur s_P. Il s'ensuit que, pour les critères du type considéré, les conclusions énoncées à la fin de ce 4.2.3 b) à propos de $c(a,b)$ et des PAMC d'ELECTRE IS et ELECTRE III demeurent valides pour $s_P(a,b)$ et pour la PAMC de PROMÉTHÉE I et II mais avec une importante réserve : tout changement de mesurage qui n'est pas de forme affine est incompatible avec le maintien de seuils constants.

Considérons maintenant le cinquième type de critère appelé *critère de niveau*. Il repose encore sur les deux seuils constants q_j et p_j qui servent à attribuer respectivement les valeurs 0 et 1 à l'indice de crédibilité partielle toujours dans les conditions de la formule (4.21), le seul changement provenant de ce que :

$$s_P(a,b) = 1/2 \quad si \; q_j < g_j(a) - g_j(b) < p_j$$

Le remplacement d'un critère g_i relevant de ce cinquième type par le critère $h_i = f\ (g_i)$ laisse s_{P_j} invariant si f_i est la forme affine après adaptation des deux seuils qui, dans ces conditions, restent constants. Avec toute autre forme de transformation monotone, cette invariance ne peut être obtenue avec le maintien de seuils constants.

Venons-en enfin au sixième et dernier type de critère appelé *"gaussien"*. Celui-ci ne fait intervenir aucun des seuils q_j et p_j. Il conduit à poser :

$$s_P(a,b) = \begin{cases} 0 & si \quad g_j(a) - g_j(b) \le 0 \\ 1 - \exp\left\{\dfrac{-(g_j(a) - g_j(b))^2}{2s^2}\right\} & si \quad g_j(a) - g_j(b) \ge 0 \end{cases}$$

où s est un paramètre dont la valeur est celle de l'écart de performance $g_j(a) - g_j(b)$ qui correspond à une crédibilité de l'assertion $a\ P_j\ b$ valant 1 – exp(-1/2) = 0.4. Cette égalité doit être respectée dans un changement de mesurage $h_j = f(g_j)$. Elle conduit notamment à changer s en as avec $f(g_j) = ag_j + b$. Le lecteur en déduira sans peine que, avec ce

sixième type de critère, s_{P_j} reste invariant dans une transformation affine de l'échelle. En revanche, avec toute autre transformation monotone, il est impossible d'attribuer une valeur au paramètre s qui garantisse l'invariance de sa signification. Pour cette raison mais aussi du fait de la forme exponentielle de l'indice de crédibilité partielle, un changement de mesurage non affine peut entraîner des variations non négligeables de cet indice.

En conclusion, **la PAMC de PROMÉTHÉE I et II est signifiante quel que soit le type de critère considéré si toutes les échelles sont de niveau intervalle. Si certaines d'entre elles sont de niveau inférieur, la perte de signifiance provient du fait que les seuils d'indifférence, de préférence, de même que le paramètre s du critère gaussien, sont des constantes.**

5. Conclusion

L'analyse de la signifiance des PAMC conduite dans les sections 3 et 4 a mis en évidence l'intérêt des différents niveaux de mesurage présentés en sections 2. Pour que ces PAMC puissent jouer le rôle qui leur est dévolu en aide à la décision, il importe non seulement que l'information disponible soit fidèlement représentée mais aussi que cette représentation (numérique) soit bien chargée du sens que présupposent les opérations arithmétiques effectuées dans le cadre de la PAMC considérée (cf. Rebai et Martel, 2001). Appliquer des outils mathématiques à un ensemble de données sans s'interroger sur certaines propriétés de ce qu'elles représentent peut facilement aboutir à des résultats non justifiés. Pour éviter de courir ce risque, il faut examiner l'impact que peut avoir sur ces résultats, le remplacement d'une échelle par une autre qui paraît tout aussi justifiée (hypothèse de travail équivalente). Nous avons montré que cela nécessite une adaptation de la valeur attribuée aux différents paramètres (poids, seuils, …) pour tenir compte de la façon dont l'information est codée.

Enfin, il nous paraît important de souligner ici le fait que le rôle dévolu de la mesure en aide à la décision et en physique n'est pas exactement le même. Dans cette seconde discipline en effet, la mesure est conçue exclusivement pour décrire, c'est-à-dire pour représenter fidèlement une réalité préexistante à la mesure. En aide à la décision, la mesure joue un rôle mais souvent pour une part seulement car elle sert aussi à **construire**. Cela tient au fait que ce qu'il s'agit de décrire n'est pas toujours ni rigoureusement, ni clairement défini (cf. Roy (1987 et 2001)). Dans bien des cas, il y a une certaine marge d'indétermination qui, pour des raisons variées (cf. notamment 4.2.3 a)), enveloppe ce que l'on cherche à appréhender. C'est pourquoi les variations qui peuvent survenir lorsqu'on passe d'une échelle à une échelle équivalente ne doivent pas être obligatoirement regardées

comme des atteintes à la signifiance mais, plus simplement, comme le reflet d'ambiguïtés inhérentes à la nature des données. Il s'ensuit que le système de préférences que fabrique une PAMC, même lorsqu'elle est utilisée dans des conditions de parfaite signifiance, ne peut qu'exceptionnellement être regardé comme une représentation fidèle d'une réalité préexistante. Ce SRP n'est, le plus souvent, qu'un point de départ, une hypothèse de travail ou un maillon d'une analyse de robustesse pouvant servir de base à une concertation ou à l'élaboration de recommandations.

Remerciements :

Les auteurs tiennent à remercier Denis Bouyssou, Pascal Lang et Vincent Mousseau qui ont bien voulu relire avec soin une première version de cet article. Les remarques et suggestions qui nous ont été transmises nous ont été fort utiles pour la mise au point de la présente version.

Références

Bana e Costa C.A., Vansnick J.C. (1994) «MACBETH – An interactive path towards the construction of cardinal value functions», *International Transactions in Operations Research* 1, 4, 489-500.

Bana e Costa C.A., Vansnick J.C. (1997), «A theoretical framework for measuring attractiveness by a categorical based evaluation technique (MACBETH)», in J. Climaco (ed.): Multicriteria Analysis, Springer, Berlin, 15-24.

Bana e Costa C.A., Vansnick J.C. (2001), «Une critique de base de l'approche de Saaty : mise en question de la méthode de la valeur propre maximale», Université Paris-Dauphine, Cahier du LAMSADE N0 175.

Barzilai J. (1998), «Measurement Foundations for Preference Function Modelling», *IEEE International Conference on Systems, Man and Cybernetics*, San Diego, California, October 1998.

Bouyssou D., Roy B. (1997), «La notion de seuils de discrimination en analyse multicritère», *Information System and Operational Research*, Vol 25, 4, 302-313.

Bouyssou, D., Marchant T., Pirlot M., Perny P., Tsoukiàs A., Vincke, Ph., (2000) Evaluation and Decision models – A critical perspective, Boston, Kluwer Acadimic Publishers, 2000.

Brans J.P., Vincke Ph. (1985), «A preference ranking organisation method: The PROMETHEE Method for Multiple Criteria Decision Making», *Management Science*, Vol. 31, 647-656.

Campbell N.R. (1938), «Symposium: Measurement and its importance for philosophy», *Proc. Arist. Soc. Suppl.*, Vol.17, 121-142. London: Harrison.

Edwards W. (1977), «How to use multiattribute utility measurement for social decisionmaking», *IEEE Transactions on Systems, Man, and Cybernetics* SMC-7,5, 326-340.

Falmagne J.C., Narens L. (1983), «Scales and Meaningfulness of qualitative laws», *Synthese*, Vol. 55, 287-325.

Hwang, C.L., Yoon K. (1981), *Multiple attribute decision making: Methods and applications – A state-of-the-art survey*, Spinger-Verlag.

Keeney R.L., Raiffa H. (1976), *Decisions with multiple objectives: Preferences and value tradeoffs*, John Wiley & Sons, New York.

Maystre L.Y., Pictet J., Simos J. (1994), *Méthodes multicritères ELECTRE*, Presse Polytechnique et Universitaire Romandes, Lausanne.

McCord, M., de Neufville, R. (1982) « Empirical demonstration that expected utility decision analysis is not operational », in Stigum and Wenstop (eds.), Foundations of utility and risk theory, D. Reidel, Dordrescht, pp. 181-199.

Rebai A., Martel J.M. (2001), « Que doit-on attendre d'une procédure d'agrégation multicritère pour des évaluations non cardinales ou mixtes », INFOR 39, 1, 51-63.

Roberts F.S. (1979), *Measurement theory*, Encyclopedia of mathematics and its applications (Vol. 7): Mathematics and the social sciences, Reading, Mass.: Addison-Wesley Publishing Co.

Roberts F.S., Franke C.H. (1976), «On the theory of uniqueness in measurement» *Journal of mathematical psychology*, Vol. 14, 211-218.

Roy B. (1969), «Classement et choix en présence de points de vue multiples : La méthode ELECTRE», *RIRO*, Vol. 2, 57-75.

Roy B. (1978), «Un algorithme de classement fondé sur une représentation floue des préférences en présence de critères multiples», *Cahier de centre d'études de recherche opérationnelle*, Vol. 20, 3-24.

Roy B. (1987), «Meaning and validity of interactive procedures as tools for decision making», *European Journal of Operational Research*, Vol 31, 297-303.

Roy, B. (1993) «Decision science or decision-aid science ?», *European Journal of Operational Research*, Vol. 66(2), 184-203.

Roy B. (1999), «Decision-aiding today: What should we expect?», in Tomas Gal, Theodor J. Stewart, Thomas Hanne (eds.), *Multicriteria Decision Making – Advances in MCDM Models, Algorithms, Theory, and Applications*, Kluwer Academic Publishers, Boston, 1-1–1.35.

Roy B. (2001), «L'aide à la decision aujourd'hui : que devrait-on en attendre ?», dans Albert David, Armand Hatchuel, Romain Laufer (eds.), les nouvelles foundations des sciences de gestion. Éditions Vuibert. Collection FNEGE, pp. 145-179.

Roy B., Bouyssou D (1993), *Aide Multicritère à la Décision : Méthodes et Cas*, Collection Gestion, Edition Economica. Paris.

Roy B., Mousseau V. (1996), «A Theoretical frameworks for analysis the notion of relative importance of criteria», Journal of Multi-Criteria Decision Analysis, Vol. 5, 145-159.

Russel B. (1938), *Principles of Mathematics*, Second edition. New York: Norton.

Saaty T.L. (1977), «A scaling method for priorities in hierarchical structures», *Journal of Mathematical Psychology*, Vol. 15, 234-281.

Saaty T.L. (1984), «Décider face à la complexité – Une approche analytique multicritère d'aide à la décision», EME, Paris.

Stevens S.S. (1946), «On the theory of scales of measurement», *Science*, Vol. 103, 677-680.

Stevens S.S. (1951), «Mathematics, measurement and psychophisics», *Handbook of experimental psychology*, John Wiley & Sons, New York, 1-49.

Suppes P. (1959), «Measurement empirical meaningfulness and three-valued logic», in Churchman and Ratoosh (eds.), Measurement : Definition and Theories, John Wiley & Sons, New York, 129-143.

Suppes, P., Zinnes J. (1963), «Basic measurement theory», in *Handbook of mathematical psychology*, Luce, Bush & Galanter (Eds), Vol. 1, 1-76, John Wiley & Sons, New York.

Vansnick J.C. (1991), «Measurement theory and decision aid», in Readings in Multiple Criteria Decision Aid, Bana e costa (Ed.), 81-100.

Vincke Ph. (1989), «L'aide multicritère à la décision», Editions Ellipses (France) et Editions de l'Université de Bruxelles (Belgique).

Von Winterfield, D., Edwards, W. (1986), «Decision Analysis and Behavioral Research», Cambridge University Press, New York.