

REVUE DE STATISTIQUE APPLIQUÉE

A. QANNARI

E. M. QANNARI

P. COURCOUX

Analyse d'un tableau ternaire d'évaluations sensorielles

Revue de statistique appliquée, tome 40, n° 4 (1992), p. 45-58

http://www.numdam.org/item?id=RSA_1992__40_4_45_0

© Société française de statistique, 1992, tous droits réservés.

L'accès aux archives de la revue « Revue de statistique appliquée » (<http://www.sfds.asso.fr/publicat/rsa.htm>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

ANALYSE D'UN TABLEAU TERNAIRE D'ÉVALUATIONS SENSORIELLES

A. Qannari, E.M. Qannari, P. Courcoux

*Laboratoire de Statistique et Informatique E.N.I.T.I.A.A. La géraudière
44072 Nantes Cedex 03*

RÉSUMÉ

Nous présentons une démarche pour analyser un ensemble de tableaux dans le contexte de l'évaluation sensorielle. Pour étudier la performance des juges, nous introduisons une métrique dont l'optimalité est étayée par le modèle à effet fixe. Cette métrique a la propriété de rendre cohérentes entre elles les évaluations des juges. Une caractérisation des produits est donnée par un compromis des tableaux.

Mots-clés : *Analyse sensorielle, Analyse en Composantes Principales, Modèle à effet fixe.*

Introduction

En évaluation sensorielle, les juges sont utilisés comme «instruments» de mesure pour caractériser les produits alimentaires sur un plan organoleptique. Ceci permet de détecter d'éventuelles variations dans la fabrication, d'aider à la conception de nouveaux produits...

L'évaluation sensorielle diffère dans ses objectifs d'une étude de consommateurs. Son rôle consiste essentiellement à établir une caractérisation objective des produits; de la même manière que l'on peut établir, par d'autres procédés, une caractérisation physico-chimique ou nutritionnelle.

Des juges sont sélectionnés sur la base de leurs aptitudes sensorielles et leurs aptitudes à exprimer objectivement les sensations perçues. Selon les objectifs visés par l'entreprise, diverses procédures (épreuves sensorielles) peuvent être mises en oeuvre. Une des épreuves sensorielles consiste à caractériser les produits par des notations sur des échelles pour un ensemble de descripteurs (variables). Un entraînement préalable des juges est nécessaire afin d'harmoniser les notations et la définition des descripteurs retenus.

Les données recueillies à l'issue de l'épreuve peuvent être présentées sous forme de tableaux (produits $\times$ descripteurs), chacun des tableaux étant associé à un juge.

Les objectifs de l'analyse de plusieurs tableaux énumérés par Escoufier (Escoufier, 1985) trouvent une signification concrète dans notre contexte :

- l'utilisateur est intéressé par une analyse globale (comparaison des tableaux) qui lui permet de porter des appréciations sur les juges et orienter, le cas échéant, les entraînements en conséquence ;
- l'utilisateur en analyse sensorielle, est intéressé par la recherche d'un compromis (au sens de résumé global) qui constitue une synthèse des évaluations individuelles des juges ;
- une analyse fine (comparaison des lignes et des colonnes des tableaux) permet de situer les difficultés des juges et de mettre en évidence des produits « instables » (grande variation dans les évaluations des juges)...

Dans la démarche exposée ici, nous comparons les tableaux, au niveau global, par l'introduction d'une métrique dont le choix est étayé par le modèle à effet fixe étudié par Caussinus (Caussinus, 1986). De cette analyse globale, nous déduisons un tableau compromis dont nous visualisons les tendances centrales par une Analyse en Composantes Principales (ACP). Pour permettre une analyse fine des tableaux, nous proposons de les projeter en éléments supplémentaires sur les espaces factoriels issus de l'ACP du compromis.

1. Position du problème

Supposons que m juges évaluent n produits selon un ensemble de p descripteurs. Il en résulte m tableaux Y_i ($i = 1, \dots, m$) de dimension $n \times p$ représentant, chacun, la configuration associée à chaque juge. Du fait des séances d'entraînement préalables, les m configurations sont sensées être identiques. En pratique, on constate quelques « dérives » que l'on peut classer comme suit :

- i) un effet différentiel dû à chaque juge : effet «juge» qui peut être évalué par une analyse de la variance ;
- ii) un facteur d'échelle propre à chaque juge : les juges n'utilisent pas la même échelle de notations en ce sens que certains d'entre eux auraient tendance à utiliser toute l'échelle alors que d'autres n'en utiliseraient qu'une partie restreinte ;
- iii) une certaine confusion entre les descripteurs : des descripteurs qui présentent une redondance entraînent, pour les juges, une confusion pouvant se traduire par le fait que, abstraction faite du facteur d'échelle, la configuration d'un juge est conforme à celle d'un autre à une rotation près.

On comprend dès lors pourquoi les méthodes procustéennes sont utilisées et adaptées à ce genre de problème (Arnold et al, 1986 ; Langron, 1981 ; Tomassone et al, 1977). Glaçon (Glaçon, 1981) énumère, en les comparant, d'autres méthodes qui pourraient trouver un champ d'application en analyse sensorielle. En particulier la méthode STATIS (Lavit, 1988) a été utilisée pour étudier ce type de données par Gonzalez (Gonzalez, 1982).

2. Traitements préliminaires des tableaux

2.1 Centrage des tableaux

Cette pratique est commune à la plupart des méthodes traitant des tableaux ternaires. On procède à un centrage de chacune des colonnes des tableaux. Ceci permet de les ramener à une même origine en éliminant l'effet différentiel dû à chacun des juges.

2.2 Facteur d'échelle

En présence d'un seul descripteur Y , nous pouvons définir les facteurs d'échelle associés aux diverses configurations de la manière suivante :

Soit $V_i(Y)$ ($i = 1, \dots, m$) la variance des notes attribuées par le i^{e} juge aux différents produits et $V(Y)$ la moyenne de ces variances, le facteur d'échelle associé au juge i peut être choisi comme étant la racine carrée de la quantité :

$$\lambda_i = \frac{V_i(Y)}{V(Y)}$$

Cette procédure généralisée à plusieurs descripteurs conduit à définir les facteurs d'échelle à partir de :

$$\lambda_i = \frac{\text{Trace}(V_i V)}{\|V\|^2}$$

où V_i est la matrice de covariance calculée à partir du tableau Y_i ($i = 1, \dots, m$) et V la moyenne de ces matrices. La norme $\|\cdot\|$ est la norme trace dans l'espace des matrices $p \times p$.

Il est aisé de vérifier que les λ_i ainsi définis réalisent le minimum de la fonction :

$$f(\lambda_1, \dots, \lambda_m) = \sum_{i=1}^m \|V_i - \lambda_i V\|^2$$

et qu'ils sont positifs.

Dans la suite, nous considérons les configurations $\tilde{X}_i$ ($i = 1, \dots, m$) « corrigées » à l'aide des facteurs d'échelle :

$$\tilde{X}_i = \frac{Y_i}{\sqrt{\lambda_i}}$$

3. Analyse globale

L'objectif de l'analyse globale est de comparer les tableaux $\tilde{X}_i$ ($i = 1, \dots, m$). Ceci signifie, dans le contexte étudié, que nous désirons comparer les performances des juges. Dans la méthode STATIS, l'analyse globale (inter-structure) est basée sur la comparaison des opérateurs associés aux tableaux (matrices de covariances,...). Dans le présent article, nous comparons les tableaux directement entre eux. Cette idée a été exploitée par Jaffrenou (Jaffrenou, 1978). La comparaison des tableaux peut se faire, en effet, sur la base de la distance euclidienne usuelle :

$$d(\tilde{X}_i, \tilde{X}_1) = \{\text{Trace}[\tilde{X}_i - \tilde{X}_1)(\tilde{X}_i - \tilde{X}_1)]\}^{1/2}; i, 1 = 1, \dots, m.$$

Il semble, néanmoins, que cette distance ne tienne pas compte de la structure d'ensemble des données (corrélations entre les variables; ...), c'est pourquoi nous introduisons une métrique qui semble plus appropriée. Le choix de cette métrique est justifié par le modèle à effet fixe pour l'ACP et par une propriété stipulant qu'elle rend les évaluations des différents juges cohérentes entre elles. Les hypothèses que nous présupposons sont conformes au protocole expérimental en analyse sensorielle.

3.1 Hypothèses de base

Les configurations associées aux différents juges sont considérées comme étant des réalisations de vecteurs aléatoires X_i ($i = 1, \dots, m$), de R^{np} vérifiant les hypothèses suivantes :

- i) Les X_i ($i = 1, \dots, m$) sont indépendants.
- ii) Les X_i ($i = 1, \dots, m$) ont la même matrice de covariance Γ ($np \times np$)

L'hypothèse i) est justifiée, dans le contexte étudié, par le fait que les évaluations des juges se font de façon non concertée. Ceci constitue un impératif formel pour une évaluation objective.

L'hypothèse ii) peut être justifiée par le fait que les juges ont suivi un même entraînement préalable et que de ce fait, ils sont «sensés» reproduire la même matrice de covariance. Nous supposons, de plus, que Γ peut s'écrire sous la forme d'une matrice diagonale par blocs :

$$\Gamma = \begin{bmatrix} \boxed{\gamma_1} & \dots & 0 & \dots \\ \dots & \boxed{\gamma_2} & 0 & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \boxed{\gamma_n} \end{bmatrix}$$

où γ_k ($k = 1, \dots, n$) est la matrice de covariance des descripteurs pour le k ième produit.

Explicitement, désignons par x_{ijk} (respectivement $x_{ij'k'}$) la variable représentant la notation affectée par le juge i au produit k (resp. k') selon le critère j (resp. j'). La covariance entre x_{ijk} et $x_{ij'k'}$ est supposée égale à :

$$\text{cov}(x_{ijk}, x_{ij'k'}) = \begin{cases} \sigma_{jj'k} & \text{si } k = k' \\ 0 & \text{sinon} \end{cases}$$

Pour bien préciser cette idée, nous allons l'illustrer par un exemple simple.

Supposons que deux produits P_1 et P_2 sont évalués selon deux descripteurs qui sont l'acidité et l'amertume.

Le vecteur associé à chaque configuration individuelle peut être représenté symboliquement de la manière suivante :

$$\begin{bmatrix} \text{Amertume de } P_1 \\ \text{Acidité de } P_1 \\ \text{Amertume de } P_2 \\ \text{Acidité de } P_2 \end{bmatrix}$$

A cette représentation peut être associée la matrice de covariance :

$$\Gamma = \begin{bmatrix} \sigma_{111} & \sigma_{121} & 0 & 0 \\ \sigma_{211} & \sigma_{221} & 0 & 0 \\ 0 & 0 & \sigma_{112} & \sigma_{122} \\ 0 & 0 & \sigma_{212} & \sigma_{222} \end{bmatrix}$$

Cette forme de la matrice de covariance traduit l'idée que l'évaluation d'un produit n'a pas d'incidence sur l'évaluation des autres ce qui constitue une hypothèse de base dans l'évaluation sensorielle.

3.2 Estimation de la matrice de covariance

Pour deux descripteurs j et j' ($j, j' = 1, \dots, p$) et un produit k ($k = 1, \dots, n$), nous disposons de m couples de valeurs qui sont les notes des m juges accordées au produit k selon les deux descripteurs. Nous nous proposons d'estimer l'élément $\sigma_{jj'k}$ à partir de la covariance (empirique) de ces couples de valeurs.

3.3 Analyse globale des tableaux

Nous nous proposons de représenter les tableaux dans un sous espace de faible dimension. Pour cela, nous réalisons une ACP du tableau Z ayant tX_i pour i^{e} ligne.

Si nous désignons par q la dimension «optimale» de représentation des X_i alors l'espace de représentation, F_q , contient tous les vecteurs X_i à l'exclusion

d'un «bruit» :

$$\forall i = 1, \dots, m \quad E(x_i) = x_i \in F_q;$$

$E(\cdot)$ désignant l'espérance mathématique.

Les hypothèses postulées au §3.1 correspondent aux hypothèses du modèle à effet fixe pour l'ACP étudié par Caussinus (Caussinus, 1986). Il en découle que le choix de la métrique $M = \Gamma^{-1}$ assure à l'espace de représentation, F_q , une optimalité analogue à celle du théorème de Gauss Markov pour le modèle linéaire. Il est démontré (Besse, 1986) que les estimateurs obtenus (l'espace F_q de représentation et les éléments x_i) sont les meilleurs (au sens d'un critère des moindres carrés) parmi tous les choix possibles de la métrique M .

En résumé, l'espace de représentation des vecteurs X_i ($i = 1, \dots, m$) est l'espace engendré par les q premiers vecteurs principaux de l'ACP du tableau Z avec le choix de la métrique :

$$\Gamma^{-1} = \begin{bmatrix} \boxed{\gamma_1^{-1}} & \dots & 0 & \dots \\ \dots & \boxed{\gamma_2^{-1}} & 0 & \dots \\ \dots & \dots & 0 & \dots \\ \dots & \dots & \dots & \boxed{\gamma_n^{-1}} \end{bmatrix}$$

Remarque :

L'objet de cette remarque est d'appuyer le choix de la métrique $M = \Gamma^{-1}$ en énonçant une propriété qui peut être utile lors de l'interprétation des résultats.

Nous nous proposons de chercher une métrique, M_0 , qui minimise l'inertie du nuage des points du tableau Z. En d'autres termes, M_0 rend cohérentes entre elles les évaluations des juges. Pour cela posons :

$$f(M) = E \left[\sum_{i=1}^m \|X_i - E(X_i)\|_M^2 \right].$$

Nous cherchons M_0 telle que $f(M_0)$ soit minimum sous la contrainte que $\det(M_0) = 1$, $\det(M_0)$ désignant le déterminant de M_0 . Compte tenu du fait que les X_i ($i = 1, \dots, m$) ont la même matrice de covariance Γ , la solution de ce problème (Romeder, 1973) est obtenue par dérivation matricielle :

$$M_0 = \left| \frac{1}{\det(\Gamma)} \right|^{-1/np} \Gamma^{-1}$$

4. Recherche d'un compromis

Le premier vecteur principal, C , de l'ACP du triplet (Z, M, I_m) , où I_m est la matrice identité d'ordre m , constitue un résumé des m vecteurs X_i ($i = 1, \dots, m$) et s'écrit sous la forme :

$$C = \sum_{i=1}^m \alpha_i X_i$$

où $\alpha = {}^t(\alpha_1, \dots, \alpha_m)$ est un vecteur propre unitaire de ZM^tZ correspondant à la plus grande valeur propre. C est un vecteur de R^{np} proche des X_i en ce sens qu'il est l'optimum de :

$$\min_{\alpha} \sum_{l=1}^m \left\| X_l - \sum_{i=1}^m \alpha_i X_i \right\|_M^2$$

Par conséquent, nous proposons de prendre comme compromis des configurations des différents juges le tableau U ($n \times p$) reconstitué à partir du vecteur C .

Le compromis ainsi obtenu permet d'obtenir une caractérisation synthétique des produits. Une classification à partir de ce tableau permet de définir des classes de produits homogènes. Une ACP de ce tableau permet de visualiser les proximités entre les produits. L'analyse fine discutée ci-après peut aider à interpréter et à définir les variables du compromis.

5. Analyse fine

L'objectif de cette étape est de permettre une comparaison des lignes et des colonnes des différents tableaux. La démarche que nous préconisons consiste à projeter en éléments supplémentaires les lignes et les colonnes des tableaux des différents juges sur les espaces propres appropriés de l'ACP du compromis. Préalablement à cette projection, nous appliquons au tableau compromis un facteur d'échelle de sorte que sa norme pour la métrique euclidienne M soit égale à la moyenne des normes des tableaux analysés.

6. Application

6.1 Les données

La démarche est illustrée par un exemple volontairement simple choisi pour son caractère pédagogique. Cinq juges A,B,C,D et E ont évalué quatre produits r , s , x et y selon trois descripteurs qui sont l'acidité, l'astringence et l'amertume. Les données sont reproduites dans le tableau 1 où, pour chacun des juges, les lignes correspondent aux produits :

TABLEAU 1
Les évaluations des cinq juges

Juge A		
7.10	5.80	6.77
7.00	4.10	6.27
2.80	4.00	3.10
3.40	6.30	4.12

Juge B		
5.80	7.00	6.10
4.40	6.80	5.00
4.00	3.40	3.85
6.10	2.60	5.22

Juge E		
4.50	7.40	5.22
6.60	5.70	6.37
2.40	5.40	3.15
3.80	3.80	3.80

Juge C		
4.70	3.90	4.50
5.20	1.90	4.37
0.60	2.10	0.97
0.80	3.70	1.52

Juge D		
6.00	5.60	5.90
6.00	4.20	5.55
4.20	4.60	4.30
3.90	5.40	4.27

Nous avons procédé à un centrage des tableaux et à une réduction par les facteurs d'échelle. Ces facteurs sont consignés dans le tableau n°2 :

TABLEAU 2
Facteurs d'échelle

Juges	Facteurs d'échelles
A	0.45626
B	0.13031
C	0.56317
D	0.11636
E	0.33343

Dans la suite, les analyses porteront sur les tableaux « corrigés » qui seront les données de base. Ces données sont regroupées dans le tableau n° 3 :

TABLEAU 3
Tableaux des données corrigées

Juge A			Juge B		
2.88	1.07	2.42	2.01	5.68	2.93
2.73	-1.35	1.71	-1.87	5.12	-0.12
-3.23	-1.49	-2.79	-2.98	-4.29	-3.30
-2.38	1.78	-1.34	-2.38	1.78	-1.34

Juge E		
0.30	3.16	1.01
3.94	0.22	3.00
-3.33	-0.30	-2.57
-0.91	-3.07	-1.45

Juge C			Juge D		
2.50	1.33	2.21	2.86	1.91	2.62
3.16	-1.33	2.04	2.86	-2.20	1.60
-2.96	-1.07	-2.49	-2.42	-1.03	-2.07
-2.70	1.07	-1.76	-3.30	1.32	-2.15

6.2 Analyse Globale

La figure n°1 donne le premier plan factoriel de l'ACP du triplet (Z, M, I_5) . Elle permet de visualiser les proximités entre les juges. La disposition des colonnes, donnée sur la figure n°2 aide à interpréter ces proximités. Dans cette figure, les chiffres se réfèrent aux numéros des descripteurs et les lettres aux produits.

Sur le premier axe le juge B s'oppose aux juges C et E. Le juge B a surévalué le produit y pour le premier descripteur et sous évalué le produit x pour le deuxième descripteur ainsi que le produit s pour le premier descripteur... les juges C et E ont eu une attitude opposée. Le second axe principal met en évidence la note particulièrement faible qu'attribue le juge A au produit x pour le premier descripteur ainsi que la note faible attribuée par le juge D au produit y pour le troisième descripteur.

6.3 Le compromis

L'ACP du triplet (Z, M, I_5) donne le compromis consigné dans le tableau n°4 :

TABLEAU 4
compromis

	V1	V2	V3
Produits <i>r</i>	4.61	6.02	4.97
Produits <i>s</i>	4.75	0.41	3.67
Produits <i>x</i>	-6.68	-3.69	-5.92
Produits <i>y</i>	-2.68	-2.75	-2.71

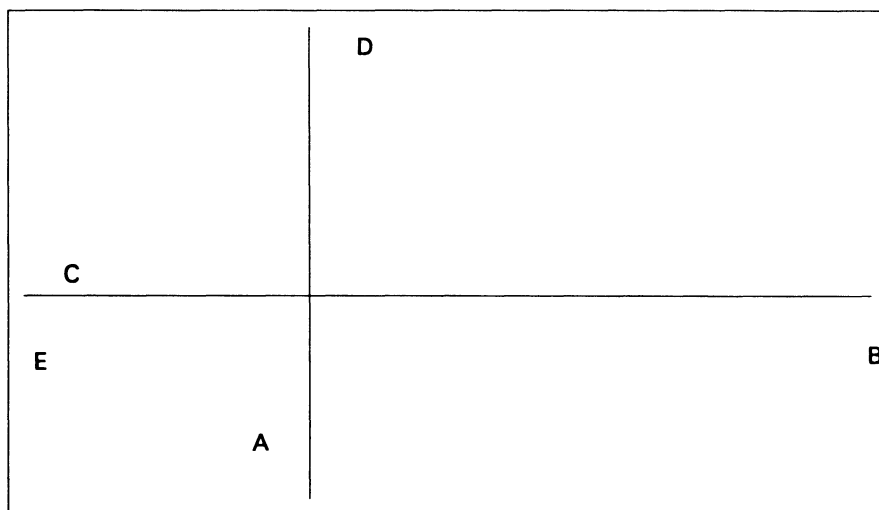


FIGURE 1

Premier plan factoriel de l'ACP du tableau Z (représentation des juges)

Le compromis constitue une synthèse des évaluations des juges. Une ACP effectuée sur ce tableau permet de positionner les produits les uns par rapport aux autres (figure n°3). Pour aider à l'interprétation, nous avons projeté sur le premier plan factoriel relatif aux variables les colonnes des tableaux des juges (figure n°4) : dans cette représentation, les lettres se réfèrent aux juges et les chiffres aux numéros des descripteurs (ainsi 1 correspond à l'acidité, 2 à l'astringence et 3 à l'amertume). Le premier axe (95 % d'inertie) met en évidence la corrélation entre l'amertume et l'acidité, il détermine un gradient d'amertume et d'acidité. Les produits s'échelonnent sur cet axe par intensité décroissante (de gauche à droite) relativement à ces descripteurs. Ainsi, les produits *r* et *s* s'opposent aux produits *x* et *y* ; les premiers ayant été jugés relativement amers. Le produit *r* se nuance du produit *s* en étant relativement plus astringent.

6.4 Analyse fine

Sur les plans factoriels de l'ACP du compromis, nous proposons de projeter en éléments supplémentaires les lignes des tableaux des 5 juges (figure n°5) de la

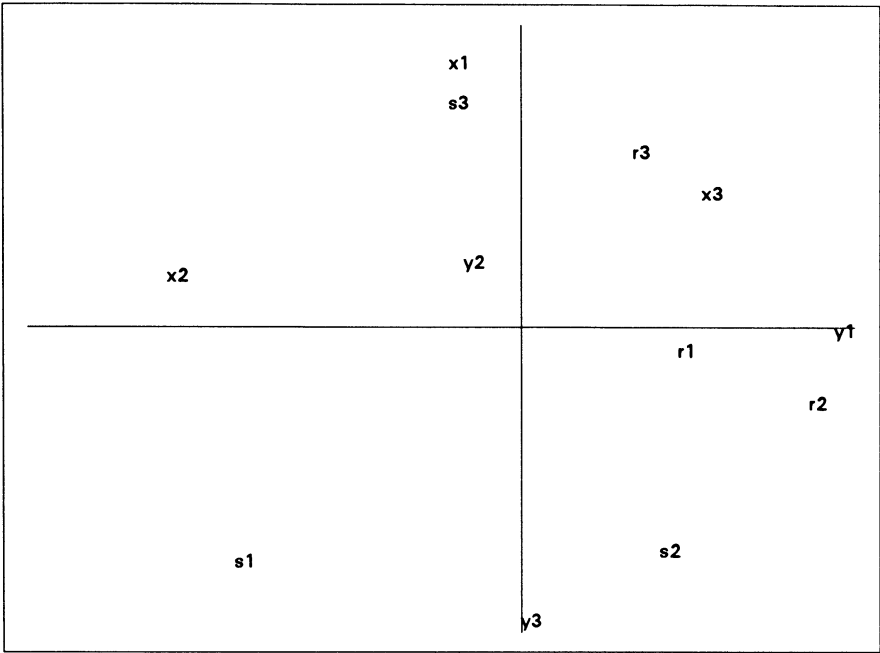


FIGURE 2

Représentation des colonnes du tableau Z dans le 1er plan factoriel (r_1 , r_2 et r_3 désignent respectivement les perceptions relatives au produit r selon les descripteurs acidité (1), astringence (2) et amertume (3); idem pour les autres produits)

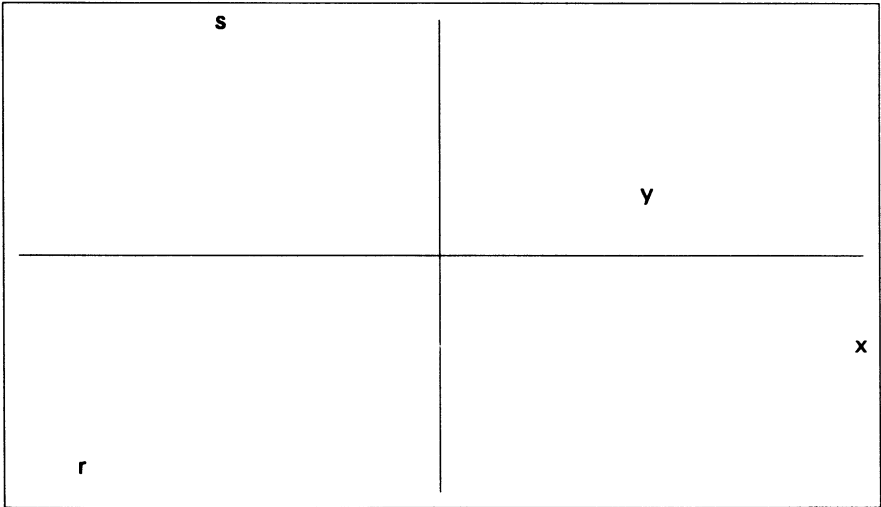


FIGURE 3

Représentation des produits sur le premier plan factoriel de l'ACP du compromis.

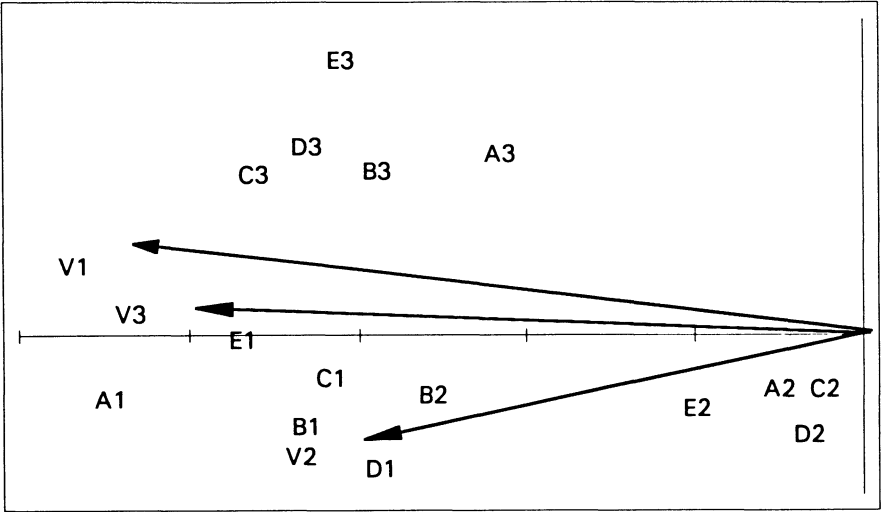


FIGURE 4

Projection des colonnes des tableaux sur le premier plan factoriel de l'ACP du compromis (Les lettres A, B, C, D et E correspondent aux juges, les chiffres aux variables et V1, V2 et V3 correspondent aux colonnes du compromis).

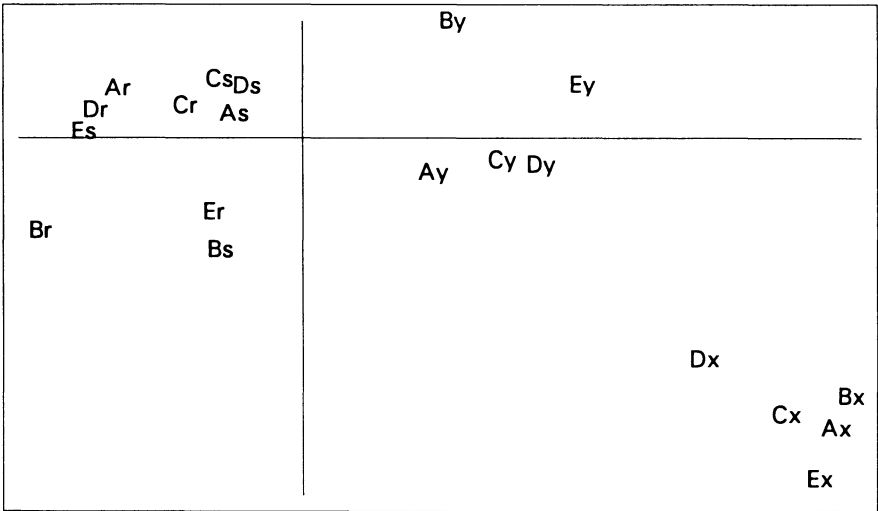


FIGURE 5

Projection des lignes des tableaux sur le premier plan de l'ACP du compromis (Ar, As, Ax et Ay désignant les évaluations des produits r, s, x et y par le juge A, idem pour les autres juges).

même manière que nous avons projeté en éléments supplémentaires les colonnes des différents tableaux sur le premier plan des variables (figure n°4).

Sur le plan des variables, il apparaît que l'astringence (descripteur n°2) a été relativement mal évaluée par le juge B. Sur le plan des individus, nous constatons que la métrique choisie a accordé les évaluations des différents juges. Nous remarquons, néanmoins, une inversion entre les produits *r* et *s* pour le juge E (Es se trouvant plus proche du «groupe» du produit *r*). Le juge B a surévalué l'astringence des produits *r* et *s* par comparaison aux autres juges.

Conclusion

La démarche exposée ici est cohérente avec les objectifs visés par l'analyse d'un tableau ternaire en analyse sensorielle. Les hypothèses énoncées sont en effet conformes au protocole d'évaluation sensorielle tel qu'il a été décrit. Il est clair que cette démarche peut être adaptée à d'autres situations impliquant le traitement de tableaux ternaires.

Bibliographie

- ARNOLD G.M., WILLIAMS A. (1986) "The Use of Generalised Procustes Techniques In Sensory Analysis" in *Statistical Procedures in Food Research*, J.R.Piggott ed., 233-253. Elsevier Applied Science. London.
- BESSE P., CAUSSINUS H., FERRE L., FINE J. (1986) "Some Guidelines for PCA". in *COMPSTAT*, Physica Verlag, Heidelberg for IASC, 23-30.
- CAUSSINUS H. (1986) «Quelques réflexions sur la part des modèles probabilistes en analyse des données». in *Data Analysis and Informatics*, IV, E. Diday et coll. éd., North Holland, Amsterdam, 151-165.
- ESCOUFIER Y. (1985) «Analyse conjointe de plusieurs tableaux de données». Numéro spécial de *Statistique et Analyse des Données* n°10 (1).
- GLACON F. (1981) «Analyse conjointe de plusieurs matrices de données : comparaison de différentes méthodes». Thèse 3^e cycle, U.S.M. Grenoble.
- GONZALEZ P.L. (1982) «Analyse statistique des données psycho-sensorielles». Thèse 3^e cycle, U.S.T.L.
- JAFFRENOU J.M. (1978) «Sur l'analyse des familles finies de variables vectorielles». Thèse, Université Claude Bernard, Lyon.
- LAFOSSE R. (1985) «Analyses procustéennes de deux tableaux». Thèse 3^e cycle, U.P.S Toulouse.
- LANGRON S. (1981) "The Statistical Treatment of Sensory Analysis Data". PHD Thesis, University of Bath, UK.
- LAVIT C. (1988) «Analyse conjointe de tableaux quantitatifs : Méthode et programmes». MASSON, Paris.

ROMEDER J.M. (1973) «Méthodes et programmes d'analyse discriminante». DUNOD, Paris.

TOMASSONE R., FLANZY C. (1977) «Présentation synthétique de diverses méthodes d'analyse des données fournies par un jury de dégustateurs». Ann. de Tech. Agricoles 26 (4), 373-418.