monoethyl ether, 111-90-0; *N*-methylpyrrolidone, 872-50-4; diethylene glycol monobutyl ether, 112-34-5; dimethyl sulfone, 67-71-0; phenol, 108-95-2; diethylene glycol monophenyl ether, 104-68-7.

LITERATURE CITED

Arctander, S. *Perfume and Flavor Chemicals*; Published by the author: Montclair, 1969.
Engel, K.-H.; Tressel, R. Studies on the volatile components of two mango varieties. *J. Agric. Food Chem.* **1981**, *31*, 796–801.
Hodge, J. E. *Symp. Foods: Chem. Physiol. Flavors, Proc.* **1967**, 472.
Hunter, G. L. K.; Bucek, W. A.; Radford, T. Volatile components of canned alphonso mango. *J. Food Sci.* **1974**, *39*, 900–903.
Idstein, H.; Schreier, P. Volatile constituents from guava (Psidium guajava, L.) fruit. *J. Agric. Food Chem.* **1985**, *33*, 138–143.
Kunishi, A. T.; Seale, P. E. *Recovery* of some volatile components from mango and guava. *Tech. Poro. Rep.* **1961**, *128*.
MacLeod, A. J.; de Troconis, N. G. Volatile flavour components of guava. *Phytochemistry* **1982**, *6*, 1339–1342.
Shiota, H. Review and study on guava fruits flavor. *Koryo* **1978**, *121*, 23–30.
Stevens, K. L.; Brekke, J. E.; Stern, D. J. Volatile constituents in guava. *J. Agric. Food Chem.* **1970**, *18*, 598–599.
Waller, G. R., Feather, M. S., Eds. *The Maillard Reaction in Foods and Nutrition*; ACS Symposium Series 215; American Chemical Society: Washington, DC, 1983.
Wilson, G. W. III; Shaw, P. E. Terpene hydrocarbons from Psidium guajava. *Phytochemistry* **1978**, *17*, 1435–1436.
Winter, M.; Malet, G.; Pfeiffer, M.; Demole, E. Structure of an odorous lactone present in the essence of jasmine (Jasminum grandiflorum L.). *Helv. Chim. Acta* **1962**, 1250.
Yamaguchi, K.; Nishimura, O.; Toda, H.; Mihara, S.; Shibamoto, T. Chemical studies on tropical fruits. In *Instrumental Analysis of Foods, Recent Progress*; Charalambous, G., Inglett, G., Eds.; Academic: New York, 1983; Vol. 2, pp 93–117.

Received for review February 18, 1988. Accepted June 13, 1988.

Selection and Classification of Volatile Compounds of Apricot Using the RV Coefficient

Pascal Schlich* and Elisabeth Guichard

The RV coefficient is a measure of similarity between two sets of variables recorded from the same sample. If the number of variables in a principal-component analysis is high, the RV coefficient allows selection of a few variables without disturbing the relative location of individuals in the first sample plots. This selection is explained and improved by a classification of variables based on the RV coefficient to define proximities between variable clusters. The groups are then submitted to a principal-coordinate analysis and minimum spanning tree using the RV matrix among the groups in order to describe relations between variable cluster. This statistical approach appears to be a very useful tool for chromatographic data handling. An example is given in a study of 56 volatile compounds quantified in 18 samples of apricots. It shows that compounds are grouped according to the chemical classes.

Progress in gas chromatography allows the separation and quantification of a great number of volatile compounds in foods or beverages. Assume that n chromatograms have been processed and p compounds quantified in each. These data are customarily arranged into a $n \times p$ matrix **X**, in which the ith row contains the p observations of variables (volatile compounds) recorded on the ith individual (chromatogram). The sample can be seen geometrically as a configuration of n points in a p-dimensional space.

Today, principal-component analysis (PCA) (Morrisson, 1976) is a classical tool in food science, as shown in the bibliography of Martens and Harries (1983) and in the methodological paper of Piggot and Sharman (1986). PCA gives an orthogonal system of principal directions of the variance of this configuration. The answer is given by the first eigenvectors of the $p \times p$ covariance matrix **C** of the compounds; if the variables have been previously autoscaled, **C** becomes the correlation matrix. Each eigenvector is defined by a linear combination of the p compounds. The interpretation of a principal component amounts to the comparison of the p coefficients of the associated linear combination.

A few problems appear when p is large (for instance higher than 30). On one hand, matrix **C** cannot be loaded in the memory of some microcomputers and the length of computing time would be prohibitive. On the other hand, and this is the main problem, the interpretation of a linear combination of so many variables would certainly be tiresome and not very convincing. In fact, a few compounds only are generally heavily loaded on the first principal axes, while the other ones only bring a background noise. However, it is often difficult to distinguish between that noise and main information and to decide which are the relevant correlations between variables and principal components. It would be of great interest to have previous knowledge of the relevant variables and then perform the PCA with these compounds only. Moreover, if p is greater than n, PCA can be performed, but $p-n$ dimensions of the sample configuration space are of course unnecessary.

The RV coefficient (Escoufier, 1970, 1973) is a measure of similarity, varying from 0 to 1, between p-dimensional and q-dimensional configurations of the same sample. It can be seen as a generalized correlation coefficient between

Laboratoire de Recherches sur les Arômes, Institut National de la Recherche Agronomique (INRA), 17 Rue Sully, 21034 Dijon Cedex, France.

0021-8561/89/1437-0142$01.50/0

two sets of p and q variables observed on the same n objects. The closer to 1 the RV is, the better correlated the two sets of variables, the nearer the two sample configurations and therefore the corresponding principal sample plots. When PCA is performed on the subset of the r variables that optimize the RV coefficient with all the p compounds, principal sample plots are obtained as close as possible to those of the whole PCA.

Any statistical method of classification of individuals can be transposed to variables with the RV coefficient to define similarity between groups of variables. The selected variables can therefore be considered as the heads of clusters of well-correlated variables. Through these heads, the unselected compounds of the PCA can find an interpretation.

In this paper, an algorithm of classification has been adapted and applied to the variables. This algorithm is currently called in France Nuées Dynamiques (Diday et al., 1980).

A principal-coordinate analysis (PCO) (Gower, 1966) can then be performed on the symmetric matrix of the RV coefficients between any two clusters of compounds obtained from the above classification. The first variable cluster plot of PCO and the superimposed minimum spanning tree (MST) (Gower, 1969) give a good picture of the relationship between clusters of compounds.

All of these statistical methods will be more precisely detailed in the next section. Application and discussion of these methods in a study (described in the Experimental Section) of 56 volatile compounds quantified in 18 samples of apricot will demonstrate the usefulness of RV coefficient for chromatographic data handling.

STATISTICAL METHODS

Selection of Variables in PCA by Optimizing the RV Coefficient. The RV coefficient and its properties were first described by Escoufier (1970, 1973). It can be used to select variables and metrics in order to reduce the number of variables in a single PCA or to establish links between two PCA on the same objects (Escoufier and Robert, 1979; Bonifas et al., 1984). First applications of these methods in food science have been recently presented by Schlich et al. (1987). The present paper more particularly deals with a single set of variables and the usual identity metric. New developments in variable classification are presented.

Let $\mathbf{X}$ be an $n \times p$ data matrix, and assume that all the p variables (columns) have been centered to have means equal to 0. It is well-known that the $n \times n$ matrix $\mathbf{XX}'$ ($\mathbf{X}'$ is the $\mathbf{X}$ transpose matrix) contains all of the usual scalar products between the objects and defines within a translation or a rotation a sample configuration. In fact, the matrix $\mathbf{S_X} = \mathbf{XX}'/[\mathrm{tr}(\mathbf{XX}'\cdot\mathbf{XX}')]^{1/2}$ is used in order to characterize a sample configuration, independently of global changes of scale. The tr notation means the trace of a square matrix, i.e. the sum of its diagonal values. Let $\mathbf{Y}$ be another $n \times q$ data matrix. To have a measure of closeness of the two sample configurations, there is a need to define one distance between $\mathbf{S_X}$ and $\mathbf{S_Y}$. For $\mathbf{A} = \mathbf{XX}'$ and $\mathbf{B} = \mathbf{YY}'$, the usual scalar product on the square matrix space is $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{tr}(\mathbf{AB}')$. It defines a norm $\|\mathbf{A}\| = [\mathrm{tr}(\mathbf{AA}')]^{1/2}$, which is equal to 1 for the $\mathbf{S}$ matrices. Thus, the distance between $\mathbf{S_X}$ and $\mathbf{S_Y}$ is given by (1).

$$\|\mathbf{S_X} - \mathbf{S_Y}\| = 2[1 - [\mathrm{tr}(\mathbf{XX}'\cdot\mathbf{YY}')/[\mathrm{tr}(\mathbf{XX}'\cdot\mathbf{XX}')\cdot\mathrm{tr}(\mathbf{YY}'\cdot\mathbf{YY}')]^{1/2}]]^{1/2} \quad (1)$$

The quantity $\mathrm{tr}(\mathbf{XX}'\cdot\mathbf{YY}')/[\mathrm{tr}(\mathbf{XX}'\cdot\mathbf{XX}')\cdot\mathrm{tr}(\mathbf{YY}'\cdot\mathbf{YY}')]^{1/2}$ is called $\mathrm{RV}(\mathbf{X}, \mathbf{Y})$. It is clear from (1) that the closer the two configurations are, the higher the RV is. As trace is a commutative operation, it can be seen that

$$\mathrm{RV}(\mathbf{X}, \mathbf{Y}) = \mathrm{tr}(\mathbf{C_{XY}}\cdot\mathbf{C_{YX}})/[\mathrm{tr}(\mathbf{C_{XX}}\cdot\mathbf{C_{XX}})\cdot\mathrm{tr}(\mathbf{C_{YY}}\cdot\mathbf{C_{YY}})]^{1/2} \quad (2)$$

in which $\mathbf{C_{XX}}$, $\mathbf{C_{YY}}$, $\mathbf{C_{XY}}$, and $\mathbf{C_{YX}}$ are equal to $\mathbf{X}'\mathbf{X}/n$, $\mathbf{Y}'\mathbf{Y}/n$, $\mathbf{X}'\mathbf{Y}/n$, and $\mathbf{Y}'\mathbf{X}/n$ and are the covariance and the cross-covariance matrices of the two sets of variables or the correlation and the cross-correlation matrices if these variables have been previously autoscaled. Equation 2 allows one to underline some interesting properties of the RV coefficient: RV is in the closed interval [0, 1]. $\mathrm{RV}(\mathbf{X}, \mathbf{Y}) = 0$ if and only if $\mathbf{C_{XY}} = 0$. $\mathrm{RV}(\mathbf{X}, \mathbf{Y}) = 1$ means that all distances between objects are proportional in the two configurations. $\mathrm{RV}(\mathbf{X}, \mathbf{Y})$ appears as a squared cosine of a generalized angle between the p-dimensional space spanned by the columns of $\mathbf{X}$ and the q-dimensional space spanned by $\mathbf{Y}$.

Another vector correlation has been earlier described by Hotelling (1936), but the advantages of the RV coefficient were illustrated by Escoufier (1973).

Let $\mathbf{Y}$ be a subset of r variables of $\mathbf{X}$, denoted by $\mathbf{X}_{(r)}$, then $\mathrm{RV}(\mathbf{X}, \mathbf{X}_{(r)})$ quantifies the ability of these variables to summarize the whole information. In other words, the closer to 1 $\mathrm{RV}(\mathbf{X}, \mathbf{X}_{(r)})$ is, the better $\mathbf{X}_{(r)}$ is as a substitute for $\mathbf{X}$ to obtain identical principal components. Because the exhaustive computation of the (p^r) RV coefficients could require too much computer time, a forward stepwise selection of variables has been therefore developed (Bonifas et al., 1984). The kth variable is then introduced in order to optimize $\mathrm{RV}(\mathbf{X}, \mathbf{X}_{(k)})$, when $k - 1$ compounds have already been selected. As soon as the increase of RV becomes negligible, or when RV is close enough to 1, the introduction of variables is stopped. No statistical test of the significance of a RV value exists. In practice, since the magnitude of RV value is comparable to that of a squared correlation, we consider that an RV value of around 0.95 indicates good similarity exists between the whole and the reduced PCA.

Classification of Variables. The algorithm called Nuées Dynamiques has been largely described by Diday et al. (1980). In our approach this method works as follows: (a) We want to obtain r clusters of variables, where r is the number of selected variables by RV coefficient. (b) We allow these variables to be the initial centers. (c) We assign each variable to the cluster whose center is the most squared correlated with it. (d) A stop condition is applied: if no variable has been moved at step c to another cluster, then we keep the current clusters. (e) For each cluster, we define a new center as the first principal component of its variables, and then go to step c.

Step b allows us to reach our target, which is to obtain as much as possible selected variables as heads of clusters of well-correlated compounds. This is the reason why a hierarchical cluster analysis was not chosen.

Principal-Coordinate Analysis and Minimum Spanning Tree of Variable Clusters. We assume that the raw data are not variables recorded on individuals, but a symmetric matrix whose elements are coefficients of similarity between the objects. PCO (Gower, 1966; Piggott and Sharman, 1986) is then able to find a low-dimensional space in which location of sample reflects as well as possible the interindividual distances expressed in these coefficients of similarity.

MST (Gower, 1969), which assumes the same kind of data as PCO, is a tree, spanning all of the objects, composed of straight-line segments joining pairs of individuals with no closed loops and for which the sum of lengths of its segments is minimum.

Table I. Volatile Compounds Ranked According to the RV Selection

step		code[a]	RV[b]
	Selected Compounds		
1	unknown 3	A1	0.738
2	2-methylethyl propanoate	B1	0.823
3	1-pentnanol	C1	0.879
4	fenchone	D1	0.902
5	linalool	E1	0.926
6	hexyl acetate	F1	0.953
7	pinocamphone	G1	0.954
8	unknown 1	H1	0.960
9	heptanal	M1	0.963
10	benzaldehyde	J1	0.965
11	*trans*-hexenyl acetate	K1	0.968
12	γ-butyrolactone	L1	0.971
13	*p*-cymen-9-ol	M2	0.976
	Unselected Compounds		
14	*trans*-2-hexenol	C2	0.977
15	δ-octalactone	H2	0.978
16	6-methyl-5-hepten-2-one	M3	0.981
17	*p*-cymen-8-ol	E2	0.982
18	1-hexanol	D2	0.985
19	dihydroactinidiolide	H3	0.986
20	nonanal	B2	0.987
21	unknown 7	E3	0.988
22	γ-nonalactone	L2	0.989
23	*trans*-2-hexenal	C3	0.990
24	terpinen-4-ol	J2	0.992
25	amyl acetate	F2	0.992
26	isopinocamphone	G2	0.993
27	phenylacetaldehyde	E4	0.994
28	benzyl alcohol	J3	0.994
29	butyl acetate	H4	0.995
30	2-pentanone	G3	0.995
31	*cis*-hexenol	I1	0.995
32	α-terpineol	E5	0.995
33	γ-decalactone	L3	0.996
34	acetophenone	B3	0.996
35	camphor	G4	0.996
36	γ-hexalactone	F3	0.996
37	unknown 2	C4	0.996
38	2-heptanone	E6	0.997
39	β-ionone	L4	0.996
40	decanal	J4	0.996
41	isobutyl propanoate	E7	0.996
42	octanal	M4	0.997
43	γ-octalactone	F4	0.997
44	γ-lactone	C5	0.998
45	unknown 6	A2	0.998
46	unknown 5	C6	0.998
47	*cis*-3-hexenyl acetate	B4	0.998
48	pentanal	M5	0.998
49	2,4-hexadienal	F5	0.998
50	2,5-hexanedione	B5	0.998
51	2-pentanol	L5	0.999
52	δ-decalactone	H5	0.999
53	2-phenylethanol	J5	0.999
54	1-butanol	B6	0.999
55	hexanal	C7	0.999
56	verbenone	A3	1.000

[a] The letter indicates the cluster the compound is grouped in after variable classification. The number (increasing with the step) identifies the compound in the cluster. [b] Value of RV coefficient between selected compounds at current step and the whole step of compounds.

As PCO is a factorial design, the superimposition of MST on the first sample plot very often gives a clear idea of variations carried on the other dimensions.

Of course, in our approach the objects are the compound clusters and coefficients of similarity are the RV coefficients.

EXPERIMENTAL SECTION

Extraction, separation by GC, and identification of the flavor volatiles from the 18 samples of apricot have been described by Guichard and Souty (1988).

The data obtained have been directly transferred to a MINI6 BULL computer for statistical analysis.

PCA has been performed with use of the SPAD statistical package (Lebart and Morineau, 1985). The directives PCO and MST from GENSTAT (INRA Biometrie, 1982) have been run on a DPS8 BULL computer under a MULTICS operating system. The other programs have been written in Fortran for the MINI6 computer.

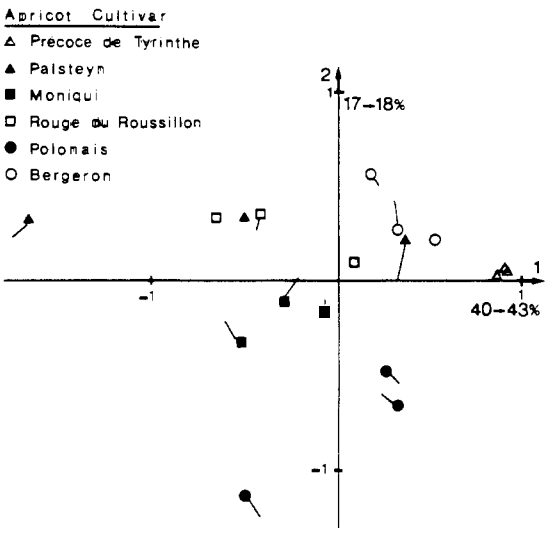

Figure 1. Superimposed sample plot (1, 2) of the whole and the reduced PCA. Ends of straight lines give new locations of apricots in the reduced PCA.

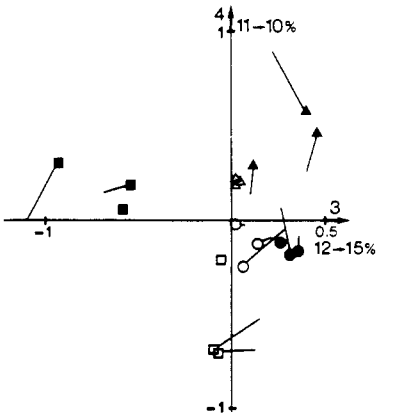

Figure 2. Superimposed sample plot (3, 4) of the whole and the reduced PCA. See caption, Figure 1.

RESULTS AND DISCUSSION

Principal-Component Analysis. Results of the whole PCA are only given in order to show the efficiency of the RV selection.

Names of the compounds are itemized in Table I, while those of apricot cultivars are first given in Table II.

The six first selected variables would certainly give a good picture of the whole PCA (RV = 0.953). However, 13 compounds (RV = 0.976) were finally retained, because this number is closer to the number of chemical classes noticed in this list of compounds.

The superimposition of the sample plot of the whole and the reduced PCA (Figures 1 and 2) shows that the location of the individuals is not significantly modified by the RV selection. The improvement of the percentage of the variance loaded on the four first principal components, given in Figures 1 and 2, is not only a consequence of the lower number of variables but also an effect of the decrease of the noise and of redundancies between variables.

Table II. Average Values (Standard Deviations) (μg/kg) of Variables Ranged in Clusters for each Cultivar

selected compounds	code	Precoce de Tyrinthe	Palsteyn	Moniqui	Rouge du Roussillon	Polonais	Bergeron
unknown 3	A1	41 (12)	1162 (677)	1246 (275)	764 (234)	481 (292)	201 (43)
unknown 6	A2	33 (11)	1319 (800)	1460 (387)	753 (222)	1054 (355)	303 (62)
verbenone	A3	141 (84)	3502 (2612)	3115 (1192)	1863 (551)	2123 (1013)	678 (226)
2-methylethyl propanoate	B1	3 (.1)	371 (177)	176 (35)	277 (76)	379 (118)	200 (64)
nonanal	B2	7 (4)	308 (113)	224 (107)	410 (94)	243 (72)	180 (46)
acetophenone	B3	11 (3)	488 (179)	218 (88)	386 (89)	348 (89)	242 (45)
cis-3-hexenyl acetate	B4	0 (0)	246 (137)	0 (0)	176 (76)	113 (49)	206 (92)
2,5-hexanedione	B5	2 (1)	142 (81)	81 (22)	106 (20)	60 (42)	53 (8)
1-butanol	B6	52 (5)	3731 (2015)	2256 (1678)	1997 (956)	957 (739)	2475 (798)
1-pentanol	C1	0 (0)	0 (0)	0 (0)	0 (0)	426 (130)	0 (0)
trans-2-hexenol	C2	137 (146)	3209 (2134)	3980 (2524)	19685 (5514)	28500 (20300)	1251 (278)
trans-2-hexenal	C3	1478 (461)	4635 (1742)	26800 (3135)	8252 (6728)	70018 (9984)	19046 (4831)
unknown 2	C4	22 (6)	154 (87)	821 (119)	333 (365)	1461 (727)	421 (142)
γ-lactone	C5	1064 (542)	6168 (1324)	3121 (630)	1673 (762)	0 (0)	3091 (948)
unknown 5	C6	0 (0)	0 (0)	0 (0)	983 (200)	2330 (594)	1408 (297)
hexanal	C7	269 (106)	1331 (502)	14049 (3918)	3513 (1506)	22828 (4626)	2845 (715)
fenchone	D1	5 (2)	93 (27)	257 (51)	266 (64)	141 (121)	82 (28)
1-hexanol	D2	62 (17)	2071 (862)	8036 (2649)	4223 (1121)	7615 (4337)	1509 (378)
linalool	E1	0 (0)	3019 (2508)	1021 (558)	864 (230)	1240 (849)	153 (33)
p-cymen-8-ol	E2	12 (7)	1582 (1305)	812 (338)	982 (200)	0 (0)	0 (0)
unknown 7	E3	80 (45)	1972 (1422)	852 (188)	621 (151)	1134 (251)	207 (79)
phenylacetaldehyde	E4	22 (7)	2405 (749)	569 (91)	217 (55)	0 (0)	341 (131)
α-terpineol	E5	21 (11)	1423 (1583)	869 (456)	635 (237)	555 (549)	290 (131)
2-heptanone	E6	3 (1)	125 (79)	0 (0)	0 (0)	0 (0)	0 (0)
isobutyl propanoate	E7	0 (0)	266 (258)	90 (12)	94 (53)	57 (44)	236 (139)
hexyl acetate	F1	1 (1)	5244 (3132)	1250 (694)	13140 (5835)	783 (680)	6953 (2270)
amyl acetate	F2	0 (0)	570 (319)	76 (51)	675 (314)	582 (114)	543 (188)
γ-hexalactone	F3	144 (40)	3164 (3914)	2856 (546)	9203 (2797)	287 (54)	3466 (279)
γ-octalactone	F4	7 (4)	1615 (817)	2022 (431)	4040 (1119)	365 (172)	1578 (444)
2,4-hexadienal	F5	0 (0)	0 (0)	0 (0)	173 (71)	0 (0)	0 (0)
pinocamphone	G1	17 (2)	357 (204)	573 (68)	231 (58)	243 (52)	77 (9)
isopinocamphone	G2	46 (16)	1793 (1145)	2185 (176)	1068 (301)	1198 (232)	332 (42)
2-pentanol	G3	7 (1)	235 (137)	318 (135)	142 (57)	314 (146)	71 (33)
camphor	G4	9 (3)	293 (199)	351 (40)	141 (44)	231 (178)	57 (9)
unknown 1	H1	12 (4)	1085 (679)	578 (122)	380 (134)	0 (0)	1643 (498)
Δ-octalactone	H2	3 (1)	171 (68)	365 (251)	470 (218)	0 (0)	404 (242)
dihydroactinidiolide	H3	408 (132)	2597 (580)	0 (0)	649 (148)	554 (214)	2708 (946)
butyl acetate	H4	0 (0)	8446 (5471)	4090 (2949)	9779 (12436)	1348 (964)	17051 (7334)
γ-decalactone	H5	563 (154)	1808 (838)	5807 (3380)	7716 (3711)	97 (63)	12629 (5255)
cis-3-hexenol	I1	58 (46)	949 (570)	473 (354)	11527 (2634)	13767 (8223)	682 (143)
benzaldehyde	J1	448 (93)	11639 (5530)	5848 (1626)	12778 (3938)	1198 (649)	2941 (448)
terpinen-4-ol	J2	46 (5)	907 (522)	1024 (175)	1713 (665)	631 (305)	343 (105)
benzyl alcohol	J3	0 (0)	824 (733)	0 (0)	1730 (356)	198 (165)	93 (18)
decanal	J4	2 (1)	149 (91)	115 (38)	154 (66)	37 (11)	27 (23)
2-phenylethanol	J5	7 (3)	128 (55)	135 (48)	210 (50)	100 (47)	19 (9)
trans-hexenyl acetate	K1	0 (0)	103 (75)	16 (12)	0 (0)	0 (0)	0 (0)
γ-butyrolactone	L1	0 (0)	0 (0)	3237 (1112)	0 (0)	50 (46)	0 (0)
γ-nonalactone	L2	0 (0)	261 (190)	680 (157)	391 (134)	154 (51)	490 (169)
γ-decalactone	L3	2112 (574)	2526 (1072)	37310 (9882)	21024 (6881)	1373 (837)	28171 (9504)
β-ionone	L4	0 (0)	0 (0)	1151 (468)	204 (47)	0 (0)	0 (0)
2-pentanol	L5	3 (2)	238 (124)	7382 (1768)	90 (36)	379 (118)	200 (64)
heptanal	M1	11 (7)	112 (48)	71 (12)	102 (21)	223 (73)	66 (22)
p-cymen-9-ol	M2	4 (2)	316 (334)	168 (58)	237 (98)	244 (87)	114 (63)
6-methyl-5-hepten-2-one	M3	6 (2)	197 (98)	138 (5)	120 (37)	166 (106)	70 (33)
octanal	M4	4 (2)	161 (79)	121 (21)	135 (23)	185 (132)	77 (26)
pentanal	M5	8 (1)	350 (300)	282 (27)	205 (62)	501 (196)	72 (16)

The factor-loading plots (Figures 3 and 4) show how cumbersome the interpretation of data with 56 variables would be. As correlations between selected variables and the principal components are almost the same in the two PCA, the almost identical sample locations have the same meaning.

The RV coefficient selects compounds among the different significant directions of the factor-loading plots and proportionally to the number of heavily loaded variables in these directions. Five variables from the thirteen (A1, B1, E1, G1, M2) are highly negatively correlated with the first principal component, which carries 40% of the variance. Two variables (F1, H1) are highly positively correlated with the second axis and one negatively (C1). M1 and J1 are more specially directed along the first and second bisectors of the first principal plane. The selection of uncorrelated variables (D1, K1, L1) with these two first principal components justifies the consideration of the third and fourth principal components (Figure 4).

Thus, interpretation of sample variations observed in Figures 1 and 2, in terms of these 13 compounds (whose quantities per variety are given in Table II), is stated below:

Precoce de Tyrinthe, positively located on the first axis, is very poor in volatile compounds.

Polonais, negatively located on the second axis, is the only cultivar containing 1-pentanol (C1). This variety does not contain the unknown 1 (H1) and is the richest in heptanal (M1).

Moniqui, located at the center of Figure 1 but clearly separated from the other varieties on the third axis, is characterized by the presence of γ-butyrolactone (L1) and

Table III. Matrix of Correlation Arranged in Blocks According to the Classification

K1	100																												
H1	12	100																											
H2	−9	53	100																										
H3	33	53	11	100																									
H4	0	72	66	36	100																								
H5	−24	57	84	37	69	100																							
F1	−5	44	71	19	66	56	100																						
F2	20	46	38	27	57	26	76	100																					
F3	−25	37	67	0	45	45	89	61	100																				
F4	−6	34	74	−3	41	47	86	56	93	100																			
F5	−20	−11	51	−16	20	30	82	43	81	81	100																		
B1	43	45	28	16	30	−3	55	75	58	60	31	100																	
B2	28	26	60	2	42	26	70	66	63	72	60	76	100																
B3	53	37	32	23	24	4	47	68	40	49	28	88	86	100															
B4	58	64	51	49	67	43	63	76	37	47	24	68	65	73	100														
B5	23	49	37	11	30	5	48	56	61	58	26	91	75	82	51	100													
B6	40	74	56	29	57	26	53	59	51	55	12	78	60	66	68	74	100												
J1	29	37	43	9	26	9	72	59	80	84	60	85	70	68	54	81	69	100											
J2	9	14	47	−19	10	12	66	57	82	88	71	72	67	56	36	64	52	87	100										
J3	−12	10	37	−14	18	9	69	49	88	85	74	63	55	37	21	61	41	85	90	100									
J4	18	28	27	−4	34	3	42	37	58	59	28	70	52	46	38	72	54	75	62	74	100								
J5	17	0	29	−24	−1	−7	51	49	69	78	60	72	61	54	29	61	45	82	95	88	71	100							
E1	35	35	−4	11	17	−27	22	52	36	26	−3	79	31	48	36	71	66	64	51	53	70	56	100						
E2	17	38	20	3	20	−11	41	41	63	56	24	80	45	49	31	84	67	85	67	77	84	68	87	100					
E3	45	34	−3	6	15	−29	18	53	29	23	−6	84	41	60	42	78	65	62	49	46	67	55	97	83	100				
E4	67	52	6	45	22	−13	16	30	15	19	−17	72	34	60	56	70	69	62	26	26	57	29	73	73	75	100			
E5	11	38	7	5	26	−9	23	47	47	35	4	67	22	28	25	63	59	60	53	63	72	56	94	86	87	58	100		
E6	60	47	−4	33	21	−25	17	39	17	12	−15	74	32	59	53	72	67	60	24	25	56	27	82	75	84	94-	64	100	
E7	1	78	39	38	43	32	43	57	52	38	2	61	25	40	45	65	69	52	40	39	38	28	66	64	62	55	70	57	100
A1	40	29	24	−12	13	−9	23	36	44	54	15	77	48	51	38	70	67	75	73	64	76	77	78	82	80	65	77	60	45
A2	37	22	8	−16	6	−13	14	40	30	37	1	74	41	48	32	64	55	58	63	52	66	70	81	72	85	56	80	55	44
A3	29	30	7	−8	13	−17	22	45	40	39	4	76	35	43	32	68	58	65	63	61	73	69	91	83	91	62	92	64	57
G1	25	18	21	−28	0	−11	6	15	30	41	0	61	39	39	15	61	54	54	62	50	65	68	65	59	69	48	67	41	34
G2	33	24	19	−21	4	−15	14	31	35	43	3	75	48	53	27	73	63	63	66	54	69	71	78	77	82	58	75	56	45
G3	28	4	3	−20	3	−25	2	34	16	22	−4	57	33	36	17	41	53	34	49	37	47	60	67	49	67	29	64	30	23
G4	40	21	8	−16	2	−19	6	32	24	30	−7	65	30	39	26	52	59	51	59	45	61	66	80	67	80	51	79	51	42
M1	20	−10	−18	−4	−4	−26	8	58	9	5	7	44	28	36	18	19	18	13	36	25	17	44	50	17	51	3	46	11	19
M2	1	35	20	4	36	0	42	65	56	41	19	77	46	48	37	73	60	59	55	63	70	55	85	79	83	47	88	57	66
M3	55	30	19	6	20	−6	29	63	33	38	9	78	48	60	56	58	69	58	64	46	58	68	82	61	83	56	75	58	51
M4	49	17	9	−3	21	−9	28	63	27	32	13	66	43	48	49	38	53	46	58	42	51	64	70	44	69	35	65	39	30
M5	23	7	−4	−13	4	−25	9	52	22	15	−1	63	33	42	21	48	45	33	48	38	46	55	79	53	80	31	76	41	42
D1	−3	−8	45	−25	17	26	42	29	54	66	53	33	44	18	18	20	27	44	70	65	51	73	22	31	17	−8	35	−18	0
D2	17	−21	−2	−40	−19	−17	−7	17	5	22	6	25	15	12	0	3	13	13	49	28	24	60	27	11	29	−11	31	−15	−7
L1	−9	−5	7	−41	−17	0	−28	−39	−5	13	−17	−5	−12	−22	−26	0	0	3	19	15	24	25	5	17	5	0	20	−16	−10
L2	−23	31	49	15	23	57	25	4	41	48	17	14	17	−2	8	19	22	24	35	39	33	29	7	25	2	3	29	−18	25
L3	−29	32	59	0	33	70	28	−3	37	50	21	−4	9	−14	12	1	12	14	29	26	20	20	−17	5	−20	−15	7	−33	12
L4	−12	−8	13	−43	−16	6	−14	−31	9	25	0	−1	−4	−18	−22	3	−4	13	30	27	29	34	3	20	4	−3	21	−18	−10
L5	−8	−2	17	−42	−12	2	−26	−38	−3	15	−19	−1	−3	−15	−25	6	14	4	20	13	24	26	6	18	7	1	18	−15	−8
C1	15	−38	−39	−19	−25	−37	−28	26	−34	−35	−17	4	−2	10	−4	−19	−18	−33	−4	−21	−22	7	13	−29	19	−28	4	−16	−17
C2	−21	−31	−6	−30	−16	−18	15	35	19	18	35	25	42	36	−5	28	−15	10	31	25	3	35	2	0	15	−23	−1	−14	−2
C3	4	−25	−22	−25	−11	−20	−29	20	−32	−29	−24	3	8	11	−5	−12	−10	−38	−6	−24	−17	3	5	−30	14	−33	2	−25	−18
C4	−7	−20	−7	−29	2	−9	−17	15	−17	−13	−12	3	18	10	−10	2	−5	−26	2	−10	−4	10	5	−18	17	−26	7	−21	−11
C5	45	45	24	67	23	23	14	4	9	18	−12	31	15	27	41	32	46	38	6	14	32	7	28	39	24	73	25	48	35
C6	−7	−2	1	5	13	15	15	55	2	−1	11	7	14	18	23	−15	−6	−21	8	−9	−24	7	−6	−36	−2	−42	−6	−32	5
C7	6	−31	−11	−41	−25	−22	−30	9	−23	−15	−16	8	13	13	−13	−5	−2	−25	13	−12	−11	21	9	−19	18	−19	6	−26	−17

I1	6	−34	1	−23	−5	−11	31	54	27	27	48	24	34	24	13	1	−1	16	48	36	8	53	17	−3	17	−29	15	−19	−7
	K1	H1	H2	H3	H4	H5	F1	F2	F3	F4	F5	B1	B2	B3	B4	B5	B6	J1	J2	J3	J4	J5	E1	E2	E3	E4	E5	E6	E7
A1	100																												
A2	94	100																											
A3	93	97	100																										
G1	93	92	86	100																									
G2	96	96	92	97	100																								
G3	73	82	75	79	80	100																							
G4	90	94	91	89	92	88	100																						
M1	31	53	50	30	40	69	56	100																					
M2	64	69	79	56	67	64	66	53	100																				
M3	80	86	84	70	80	81	90	69	71	100																			
M4	69	78	75	59	67	80	84	80	61	94	100																		
M5	62	78	78	63	72	85	83	85	80	87	85	100																	
D1	56	51	45	58	50	61	56	39	40	51	58	44	100																
D2	55	67	53	65	59	79	71	65	23	61	71	61	70	100															
L1	51	50	39	70	53	40	47	−11	0	12	11	8	48	57	100														
L2	39	34	32	49	36	28	30	−2	26	14	11	8	63	32	65	100													
L3	25	20	15	36	21	7	15	−24	0	1	1	−16	56	30	67	88	100												
L4	53	50	41	69	52	33	44	−11	2	11	13	6	55	55	98	68	72	100											
L5	52	48	36	73	57	50	50	−10	3	15	12	13	51	58	96	63	63	90	100										
C1	−6	22	12	2	7	48	27	83	16	39	53	61	13	57	−16	−29	−38	−21	−15	100									
C2	2	17	11	8	14	13	−4	43	17	7	9	26	6	26	−14	−11	−17	−8	−14	45	100								
C3	0	28	14	18	19	57	31	75	17	35	49	59	24	64	8	−3	−11	1	13	91	44	100							
C4	10	33	21	27	27	45	24	55	11	24	32	44	16	55	18	7	3	14	23	64	66	79	100						
C5	35	20	24	24	23	4	17	−22	13	19	2	−9	4	−20	14	40	21	12	15	−50	−42	−48	−36	100					
C6	−22	0	−6	−21	−14	23	2	71	14	27	43	39	18	40	−34	−9	−8	−33	−34	79	45	75	51	−49	100				
C7	18	40	23	39	37	68	47	71	17	43	52	64	39	77	28	6	−2	21	36	82	40	93	75	−44	59	100			
I1	12	25	21	6	13	41	26	23	23	47	62	54	46	59	−25	−15	−17	−18	−24	72	58	57	52	−46	73	56	100		
	A1	A2	A3	G1	G2	G3	G4	M1	M2	M3	M4	M5	D1	D2	L1	L2	L3	L4	L5	C1	C2	C3	C4	C5	C6	C7	I1		

[a] Codes of compounds refer to those of Table I. Correlations have been multiplied by 100.

Table IV. Matrix of RV Coefficients between Clusters of Compounds, Measures of Cluster Cohesion

	K	H	F	B	J	E	A	G	M	D	L	C	I
K	1.00												
H	0.05	1.00											
F	0.04	0.34	1.00										
B	0.23	0.30	0.46	1.00									
J	0.04	0.09	0.63	0.51	1.00								
E	0.21	0.16	0.18	0.54	0.45	1.00							
A	0.13	0.04	0.14	0.41	0.55	0.68	1.00						
G	0.11	0.04	0.08	0.34	0.42	0.51	0.89	1.00					
M	0.15	0.05	0.19	0.39	0.36	0.53	0.63	0.65	1.00				
D	0.02	0.11	0.19	0.07	0.36	0.07	0.36	0.50	0.40	1.00			
L	0.05	0.18	0.10	0.03	0.09	0.03	0.21	0.28	0.02	0.37	1.00		
C	0.06	0.15	0.09	0.07	0.07	0.10	0.06	0.14	0.31	0.26	0.07	1.00	
I	0.00	0.05	0.19	0.05	0.16	0.04	0.04	0.07	0.38	0.32	0.05	0.52	1.00
	K	H	F	B	J	E	A	G	M	D	L	C	I
PC1[a]	100%	64%	80%	78%	84%	79%	96%	90%	81%	85%	81%	66%	100%
MCS[b]	1.00	0.55	0.73	0.68	0.78	0.69	0.95	0.86	0.73	0.83	0.74	0.54	1.00

[a] PC1 is the percentage of variance accounted by the first PC in the PCA of the only variables of this cluster. [b] MSC is the mean squared correlation in this cluster.

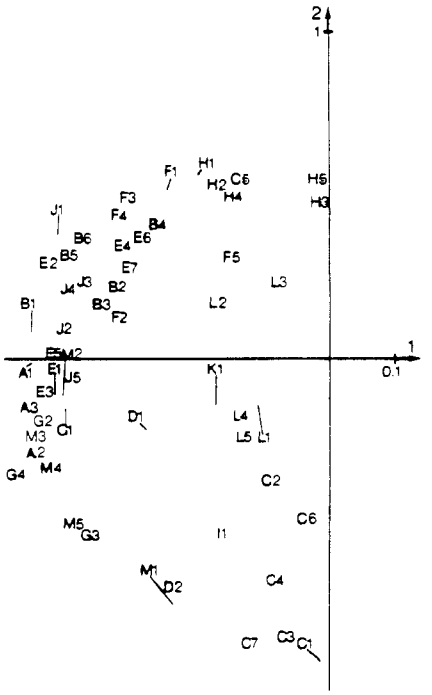

Figure 3. Correlations between volatile compounds and the two first principal components of the whole PCA. Compound codes refer to those in Table I. Ends of straight lines give new correlations of selected variables in the reduced PCA.

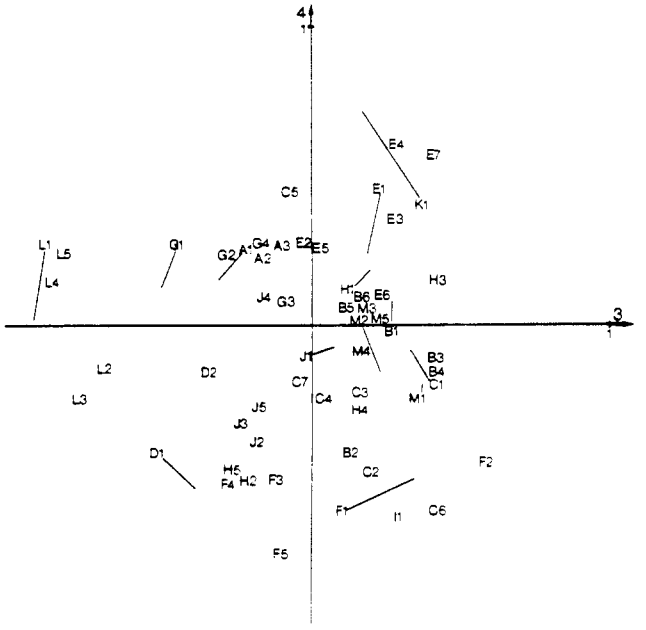

Figure 4. Correlations between volatile compounds and the third and fourth principal components of the whole PCA. See caption, Figure 3.

to lesser extent by pinocamphone (G1), unknown 3 (A1), and fenchone (D1).

Palsteyn and Rouge du Roussillon, globally located on the second bisector of Figure 1, contain the highest amounts of benzaldehyde (J1) and are also rich in hexyl acetate (F1). These two cultivars are separated along the fourth axis due to the presence of *trans*-hexenyl acetate (K1) in Palsteyn.

Bergeron, located on the first bisector of Figure 1, is rather poor in volatile compounds but the richest in unknown 1 (H1) and relatively rich in hexyl acetate (F1).

As stated in the introduction, our approach will now consist of establishing links between the selected compounds and the others.

Classification of Compounds. The algorithm of Nuées Dynamiques gives, in four iterations, the stable distribution in 13 clusters stated in Table II.

In this classification, each cluster contains at least one of the RV-selected compounds, except cluster I composed of *cis*-3-hexenol alone. These variables are truly good heads of groups because all of the information given by the above interpretation of the reduced PCA is conserved when any compound is substituted to its head of group (Table II). For instance, Polonais, which is the only variety containing 1-pentanol (C1), is the variety richest in all of the other components of cluster C and the poorest in γ-lactone (C5). This opposition is absolutely logical because two negatively correlated variables can have a high squared correlation and then be in the same group. The whole correlation matrix of Table III shows that C5 is the only example of this situation. All of the other interpretations could be checked by reference to Table II. However, Bergeron variety, which has been found to be only relatively rich in hexyl acetate (F1), does not contain 2,4-hexadienal (F5). This aldehyde has only been identified in Rouge du Roussillon. So variables, present in only one variety, are not systematically selected by RV coefficient: 1-pentanol (C1) represents the head of cluster C, but on the contrary, 2,4-hexadienal (F5) is the last selected variable in cluster F.

The standard deviations show rather large values (Table II). These variations account for the natural variability of the apricots within the same variety. In spite of this variability, the varietal interpretation of the reduced PCA has been successfully transposed to the unselected compounds. So one could ask if any selection of 13 compounds (one per cluster) would give a satisfactory RV value. This hypothesis has been tested by simulation: An average RV value equal to 0.944 (with a standard deviation equal to 0.020) has been obtained from 10 000 drawing lots according to this classification. When the simulation was carried out from 10 000 drawing lots without any classification, the average RV value was only equal to 0.907 (with a standard deviation of 0.036). These two simulations show that, with the classification, the RV value is significantly higher than without classification, accounting for a fairly good flexibility to choose variables in a same cluster but still does not reach the high RV value (0.976) obtained with the first 13 selected variables.

We now examine how compounds are distributed into the different clusters (Table II): Clusters A and G are only composed of sesquiterpenic ketones. The mass spectra of unknowns 3 and 6 of cluster A (Guichard and Souty, 1988) show that these compounds are also sesquiterpenic ketones. Cluster C is composed of C_5 or C_6 aliphatic compounds, except for γ-lactone, which is negatively correlated with the others as previously discussed. Other aliphatic compounds are located in cluster M. Cluster E is principally composed by terpenic alcohols. Cluster F contains acetates and γ-lactones. The other γ-lactones are located in cluster L. The δ-lactones are all present in cluster H. Cluster J contains phenolic compounds. Cluster B cannot be characterized by any chemical class. Clusters J, K, and D contain only one or two compounds.

It is noticable that compounds seem to aggregate globally according to the chemical classes.

Diagonal blocks of the correlation matrix, given in Table III, point out the good statistical homogeneity of the clusters. This fact is confirmed by the high mean-squared correlation (MSC) of blocks and the high percentage associated to the first principal component (PC1) in each cluster given in Table IV. Note that in Tables III and IV

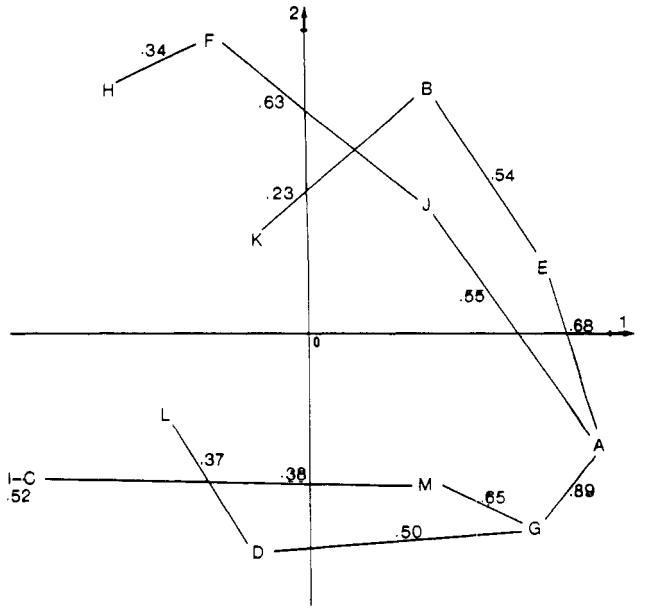

Figure 5. First PCO sample plot and superimposed MST of compound clusters. Cluster letters and RV values refer to those in Table IV.

the order of clusters has been arranged according to cluster similarities. So, in this matrix, the absolute values of correlations generally decrease as distances from the diagonal increase.

Choosing 13 clusters probably forces too sharp a classification for the whole variation. First, this probably explains the presence of small clusters. Second, some clusters could be very similar; for instance, the seven sesquiterpenic ketones of clusters A and G have the same meaning in the factor loading plots of PCA (Figure 3 and 4). All of these reasons justify the next study of cluster proximities.

Proximities between Cluster Compounds. A study of the relationship between groups amounts to consideration of the extra-diagonal blocks of Table III. A useful summary of these data is obtained by computing the RV matrix between the clusters (Table IV). The first principal-coordinate plot and the minimum spanning tree of this similarity matrix are given in Figure 5.

This tree is composed of four branches converging to cluster A and G. Axis 1, which accounts for 21% of the total variance, shows the ability of the clusters to discriminate the different cultivars. The clusters, which are typical of a variety as seen with the results of PCA, are located on the negative part of this axis and the less discriminative clusters on the positive part. Axis 2, which accounts for 18% of the total variance, gives the variations between the branches.

According to these observations, the sesquiterpenic ketones found in clusters A and G seem to be a base in the apricot aroma because no large variations in quantities are observed between the different cultivars.

We now examine the different branches. The branch E, B, K contains the terpenic alcohols (cluster E), which are known to give an important contribution to the overall apricot quality (Tang and Jennings, 1967, 1968). These compounds are more abundant in Palsteyn, and this is also true, but in a lower extent, for compounds of clusters B and K. Phenolic compounds, such as benzaldehyde (cluster J), some γ-lactones (cluster F), and the δ-lactones (cluster M) are joined in the same branch. These are key compounds for apricot aroma (Rodriguez et al., 1980; Chairotte et al., 1981). However, if clusters J and F are typical of Rouge du Roussillon, cluster M is more representative of the cultivar Bergeron. These two branches, which contain the most important compounds for apricot aroma, are located on the positive part of axis 2 and allow us to distinguish three typical varieties from the others (Guichard and Souty, 1988).

The branch M, C, I is located on the negative part of axis 2. It contains most of the aliphatic alcohols and aldehydes. These compounds have very high-level concentrations in Polonais and are not typical for apricot aroma because of their herbaceous aromatic notes. The last branch D, L with two ketones and three γ-lactones characterizes Moniqui variety, which is also rich in sesquiterpenic ketones. These latter compounds are responsible for flowery aromatic notes. Unfortunately, cluster L is very poorly represented in the first coordinate plot; in fact it defines the third coordinate, and so its compounds are uncorrelated with the other γ-lactones.

Figure 5 appears as a good chemical map of essential variability in our apricot sample, although the two first coordinates account for only 40% of the global variance between clusters.

Some varieties have been found to be the richest in compounds having typical apricot flavor notes. However, a sensory analysis would be necessary to postulate that these varieties possess the most typical overall apricot aroma.

CONCLUSION

Sampling and classification of individuals are classical topics in statistics. Allowing the separation of many compounds, gas chromatography transposes these topics from individuals to variables. This paper demonstrated the efficiency of the RV coefficient to classify the variables and reduce their number.

More generally, the RV coefficient can be considered as an unifying tool for linear multivariate statistical methods (Robert and Escoufier, 1976). It could be of a great interest to obtain correlations between sensory and instrumental data and should therefore become a powerful tool for food data analysis.

ACKNOWLEDGMENT

We are grateful to Pr. Y. Escoufier for valuable advice and to S. Fors and S. O'Keefe for help in translation.

Registry No. 2-Methylethyl propanoate, 106-36-5; 1-pentanol, 71-41-0; fenchone, 1195-79-5; linalool, 78-70-6; hexyl acetate, 142-92-7; pinocamphone, 547-60-4; heptanal, 111-71-7; benzaldehyde, 100-52-7; *trans*-hexenyl acetate, 117605-10-4; γ-butyrolactone, 96-48-0; *p*-cymen-9-ol, 4371-50-0; *trans*-2-hexenol, 928-95-0; δ-octalactone, 698-76-0; 6-methyl-5-hepten-3-one, 86883-66-1; *p*-cymen-8-ol, 1197-01-9; 1-hexanol, 111-27-3; dihydroactinidiolide, 17092-92-1; nonanal, 124-19-6; γ-nonalactone, 104-61-0; *trans*-2-hexenal, 6728-26-3; terpinen-4-ol, 562-74-3; amyl acetate, 628-63-7; isopinocamphone, 15358-88-0; phenylacetaldehyde, 122-78-1; benzyl alcohol, 100-51-6; butyl acetate, 123-86-4; 2-pentanone, 107-87-9; *cis*-3-hexenol, 928-96-1; α-terpineol, 98-55-5; γ-decalactone, 706-14-9; acetophenone, 98-86-2; camphor, 76-22-2; γ-hexalactone, 695-06-7; 2-heptanone, 110-43-0; β-ionone, 79-77-6; decanal, 112-31-2; isobutyl propanoate, 540-42-1; octanal, 124-13-0; γ-octalactone, 104-50-7; *cis*-3-hexenyl acetate, 3681-71-8; pentanal, 110-62-3; 2,4-hexadienal, 117527-48-7; 2,5-hexanedione, 110-13-4; 2-pentanal, 6032-29-7; δ-decalactone, 705-86-2; 2-phenylethanol, 60-12-8; 1-butanol, 71-36-3; hexanal, 66-25-1; verbenone, 80-57-9.

LITERATURE CITED

Bonifas, L.; Escoufier, Y.; Gonzales, P. L.; Sabatier, R. Choix de variables en analyse en composantes principales (Variable choice in principal component analysis). *Rev. Stat. Appl.* **1984**, *32*, 5–15.

Chairotte, G.; Rodriguez, F.; Crouzet, J. Characterization of additional volatile flavor components of apricot. *J. Food Sci.* **1981**, *46*, 1898–1906.

Diday, E.; et al. *Optimisation en Classification Automatique* (Optimization in classification); INRIA: Rocquencourt, France, 1980.

Escoufier, Y. Echantillonage dans une population de variables aleatoires reelles (Sampling in a population of real random variables). *Publ. Inst. Stat. Univ. Paris* **1970**, *19*, 1–47.

Escoufier, Y. Le traitement des variables vectorielles (The treatment of vectorial variables). *Biometrics* **1973**, *29*, 751–760.

Escoufier, Y.; Robert, P. Choosing variables and metrics by optimizing the RV coefficient. In *Optimizing Methods in Statistics*; Rustagi, J. S., Ed.; Academic: New York, 1979; pp 205–219.

Gower, J. C. Some distances properties of latent root and vector methods used in multivariate analysis. *Biometrika* **1966**, *53*, 325–338.

Gower, J. C. Minimum Spanning Trees and Single Linkage Cluster Analysis. *Appl. Stat.* **1969**, *18*, 54–64.

Guichard, E.; Souty, M. Comparison of the relative quantities of aroma compounds found in fresh apricot (Prunus armeniaca) from six different varieties. *Z. Lebensm. Unters. Forsch.* **1988**, *186*, 301–307.

Hotelling, H. Relations between two sets of variates. *Biometrika* **1936**, *28*, 321–377.

INRA Biometrie, Genstat un langage statistique (Genstat a statistical language), Paris, France, 1982.

Lebart, L.; Morineau, A.; et al. *SPAD*; CESIA: Paris, France; 1985.

Martens, M.; Harries, J. M. A bibliography of multivariate statistical methods in food science and technology. In *Food Research and Data Analysis*; Martens, H., Russwurm, H., Eds.; Applied Science: London, 1983; p 493.

Morrisson, D. F. *Multivariate Statistical Methods*, 2nd ed.; McGraw-Hill: New York, 1976.

Piggott, J. R.; Sharman, K. Methods to aid interpretation of multidimensional data. In *Statistical Procedures in Food Research*; Piggott, J. R., Ed.; Elsevier Applied Science: London, 1986; p 181.

Robert, P.; Escoufier, Y. A unifying tool for linear multivariate statistical methods: the RV coefficient. *Appl. Stat.* **1976**, *25*, 257–265.

Rodriguez, F.; Seck, S.; Crouzet, J. Constituants volatils de l'abricot variete Rouge du Roussillon (Volatile constituents of the apricot variety Roussillon Red). *Lebensm. Wiss. Technol.* **1980**, *13*, 152–155.

Schlich, P.; Issanchou, S.; Guichard, E.; Etievant, P.; Adda, J. RV coefficient: a new approach to select variables in PCA and to get correlations between sensory and instrumental data. In *Flavour Science and Technology*; Martens, M., Dalen, G. A., Russwurm, H., Jr., Eds.; Wiley: Chichester, 1987; p 469.

Tang, C. S.; Jennings, W. G. Volatile compounds of apricot. *J. Agric. Food Chem.* **1967**, *15*, 24–28.

Tang, C. S.; Jennings, W. G. Lactonic compounds of apricot. *J. Agric. Food Chem.* **1968**, *16*, 252–254.

Received for review October 14, 1987. Accepted June 27, 1988.