

ANÁLISE EM COMPONENTES PRINCIPAIS EXEMPLO DE APLICAÇÃO NA DIFERENCIACÃO DE MOSTO DE UVA

JEAN PIERRE ROSIER*

As análises estatísticas multidimensionais, como é o caso da análise em componentes principais (ACP), adaptam-se perfeitamente para a interpretação de resultados obtidos em enologia. Os mostos e os vinhos podem ser definidos pelo conjunto de suas características constitutivas. Pequenas diferenças não significativas em cada um dos constituintes podem fazer com que um vinho seja efetivamente distinto de outro. A ACP simplifica a compreensão de um problema complexo com o mínimo de perdas de informação. Este trabalho apresenta de forma prática a interpretação de uma ACP realizada sobre a composição em ácidos orgânicos de mostos de uva provenientes de quatro variedades e de dois solos distintos.

1 INTRODUÇÃO

As análises multidimensionais são técnicas estatísticas antigas. Segundo HOTELLING citado por PHILIPPEAU, a análise em componentes principais por exemplo, é conhecida desde 1933. Entretanto sua utilização é mais recente, foi somente após o aparecimento da informática que realmente difundiram-se estas técnicas.

Um vinho, ou um mosto podem ser definidos pelo conjunto de suas características constitutivas. A grande quantidade de substâncias que reagem entre si, fazem da interpretação de um conjunto de dados enológicos um trabalho difícil. Quando o número de amostras é elevado, as dificuldades de interpretação são tais que o auxílio da informática é indispensável.

As análises multidimensionais são um meio de abordar uma planilha de dados, considerando todas as variáveis simultaneamente e conferindo-lhes exatamente o mesmo peso.

*Eng. Agr. Doutor em Enologia, Pesquisador da EPAGRI/Estação Experimental de Videira, Santa Catarina.

É efetuado um estudo global incluindo todos os dados, com o objetivo de colocar em evidência os fenômenos mais importantes, sem hipóteses estabelecidas previamente.

Na realidade elas reduzem uma realidade complexa de um conjunto de dados, descrevendo os fatos essenciais sob forma de gráficos com mais de uma dimensão. Desta maneira, a compreensão do problema estudado é simplificada com perda mínima de informação. Logo, trata-se de ferramenta que se adapta bem ao tipo de dados correntemente utilizados em análises enológicas.

Os dados brutos são normalmente apresentados sob forma de quadros onde distinguem-se dois grupos: os indivíduos (neste caso os mostos) e os valores relativos a estes indivíduos (as variáveis medidas).

Neste trabalho são executadas sucessivamente as principais etapas de interpretação da análise em componentes principais (ACP) sem entrar nas concepções matemáticas deste método.

A análise em componentes principais é um método estatístico descritivo destinado a apresentar sob forma gráfica um quadro de dados que contém em "ativo" exclusivamente variáveis quantitativas ou qualitativas ordenadas, conservando o máximo da informação inicial.

Segundo PHILIPPEAU, o interesse na utilização da análise em componentes principais é duplo, primeiramente porque permite reduzir o número de variáveis a serem consideradas em função de suas associações, mostrando também aquelas que variam num mesmo sentido e as que se opõe. Além disso, fornece interpretação complementar das observações, sua repartição, semelhanças e diferenças (2).

Um outro aspecto do estudo das ligações entre as variáveis, segundo ESCOFIER & PAGES, consiste em resumir o conjunto das variáveis iniciais em pequeno número de variáveis sintéticas chamadas componentes principais. Logo, um componente principal pode ser considerado como a síntese de um grupo de variáveis ligadas entre si (1).

A ACP permite a utilização de elementos suplementares (variáveis ou indivíduos), que não interferem no momento dos cálculos dos eixos principais, mas que são projetados nos gráficos podendo servir de ponto de referência ou serem confrontados com indivíduos ativos no momento da interpretação.

A título de exemplo serão seguidos os principais passos de uma ACP realizada sobre os valores de dosagem dos ácidos orgânicos e das análises clássicas de mostos de uva. Os dados de base estão apresentados nos Quadros 1 a 3.

QUADRO 1 - TEORES EM ÁCIDOS ORGÂNICOS, pH, ACIDEZ TOTAL, GRAU PROVÁVEL E POTÁSSIO - MOSTOS DE 1989 - PARCELA ARENOSA

variedades	SÉMILLON				ARRILOBA					
	Porta-enxertos		confundidos		Fercal		101-14		Gravesac	
Sist.condução	Lira	tradic	Lira	tratit	Lira	tradic	Lira	tradic	Lira	tratit
Adubação	fraca	fraca	forte	forte	confundida					
Ácido tartárico*	5,20	7,00	6,40	6,80	5,15	6,32	5,87	6,23	4,47	6,08
Ácido málico*	1,70	1,80	2,00	1,90	1,07	1,17	1,33	1,43	0,95	1,37
Ácido shikímico**	10,70	9,60	10,90	10,60	4,63	3,82	5,70	3,91	5,82	5,58
Ácido cítrico**	90,00	170,00	80,00	150,00	170,00	100,00	130,00	170,00	170,00	240,00
pH	3,08	3,16	3,18	3,09	3,14	3,26	3,13	3,23	3,36	3,33
Acidez total***	4,60	4,90	4,80	4,10	3,60	3,20	3,10	3,00	3,10	3,60
Grau provável	10,30	10,00	11,20	9,80	13,10	12,80	13,50	13,20	13,00	13,60
Potássio**	741,00	1.170,00	1.209,00	1.306,00	468,00	819,00	546,00	702,00	448,00	682,00

* g/L

** mg/L

*** g/L H₂SO₄

As variáveis são o pH, a acidez total (AT), o grau provável (GP), o potássio (K) o ácido tartárico (TAR), o ácido málico (MAL), o ácido cítrico (CIT) e o ácido shikímico (SHI).

QUADRO 2 - TEORES EM ÁCIDOS ORGÂNICOS, pH, ACIDEZ TOTAL, GRAU PROVÁVEL E POTÁSSIO - MOSTOS DE 1989 - PARCELA ARGILOSA

variedades	SÉMILLON				SAUVIGNON			
	Fercal				Fercal			
Porta-enxertos	Lira	tradic	Lira	tratit	Lira	tradic	Lira	tratit
Sist.condução	fraca	fraca	forte	forte	fraca	fraca	forte	forte
Adubação								
Ácido tartárico*	6,80	5,80	6,80	7,10	7,74	7,60	7,66	8,05
Ácido málico*	2,60	2,60	3,20	3,10	4,11	4,38	3,69	5,05
Ácido shikímico**	12,90	10,20	13,70	12,10	25,55	21,83	27,95	23,86
Ácido cítrico**	270,00	350,00	330,00	380,00	100,00	180,00	150,00	230,00
pH	3,04	3,06	3,07	3,08	2,83	2,80	2,90	2,84
Acidez total***	4,00	5,10	4,50	5,10	7,50	7,80	7,60	8,30
Grau provável	10,00	10,00	9,60	9,60	11,40	11,00	12,20	10,60
Potássio**	1.131,00	823,00	1.387,00	1.095,00	858,00	994,00	1.228,00	1.253,00

* g/L

** mg/L

*** g/L H₂SO₄

As observações são em número de 26 e representam entre outros parâmetros os mostos de quatro variedades, Sémillon (SE), Sauvignon (SA), Arriloba (AR) e Liliorila (LI), repartidas sobre dois solos distintos, tipo argiloso (AC) e arenoso com cascalho (SG).

QUADRO 3 - TEORES EM ÁCIDOS ORGÂNICOS, pH, ACIDEZ TÓTAL, GRAU PROVÁVEL E POTÁSSIO - MOSTOS DE 1989 - PARCELA ARGILOSA

variedades	ARRILOBA				LILIORILA			
	Fercal		Fercal					
Porta-enxertos	Lira fraca	tradit fraca	Lira forte	tratit forte	Lira fraca	tradit fraca	Lira forte	tratit forte
Sist.condução	8,47	8,23	8,34	8,50	7,14	7,74	8,66	7,84
Adubação	2,58	2,64	2,83	2,88	3,43	3,33	4,00	3,70
Ácido tartárico*	10,80	9,30	10,95	9,10	63,56	50,92	64,60	50,85
Ácido málico*	210,00	260,00	190,00	290,00	170,00	230,00	210,00	270,00
Ácido shikímico**								
Ácido cítrico**								
pH	3,17	3,18	3,19	3,13	2,99	2,95	2,93	3,02
Acidez total***	4,90	4,30	4,60	4,60	6,40	7,00	7,00	6,50
Grau provável	13,20	12,60	12,30	12,70	12,40	12,45	12,00	12,60
Potássio**	1.053,00	1.092,00	1.365,00	1.414,00	1.287,00	799,00	1.306,00	1.174,00

* g/L **mg/L ***g/L H₂SO₄

O objetivo deste trabalho é estudar a variabilidade das amostras segundo a diferenciação dos mostos em relação às variedades e quanto aos tipos de solo.

2 ANÁLISE EM COMPONENTES PRINCIPAIS

Antes de entrar na análise multidimensional propriamente dita é interessante que se observe o comportamento individual das variedades, em relação às suas médias e desvios padrões. Esta análise preliminar permite verificar se as diferenças entre os indivíduos têm ou não significado enológico.

O Quadro 4 mostra as médias e os desvios padrões de cada variável.

QUADRO 4 - MÉDIAS E DESVIOS PADRÕES DA POPULAÇÃO

Variáveis	Médias		desvio-padrão
pH	3,08		0,14
AT	5,12	g/L H ₂ SO ₄	1,58
GP	11,73	% vol	1,32
K	982	mg/L	265
TAR	6,99	g/L	1,11
MAL	2,6	g/L	1,1
CIT	0,2	g/L	0,08
SHI	17,9	g/L	16,14

Através do exame deste Quadro constata-se que a população apresenta importantes variações no conjunto das variáveis, particularmente o desvio padrão relativo ao ácido shikímico que é da mesma ordem de grandeza que a média. Estas variações podem ser consideradas como diferenças na prática enológica.

A ligação que as variáveis podem ter entre si também é muito importante; o exame da matriz de correlações (Quadro 5) permite que se tenha uma primeira idéia da ligação entre as variáveis.

QUADRO 5 - MATRIZ DE CORRELAÇÕES

	pH	AT	GP	K	TAR	MAL	CIT	SHI
pH	1							
AT	-0,894	1						
GP	0,380	-0,234	1					
K	-0,353	0,428	-0,419	1				
TAR	-0,487	0,603	0,015	0,681	1			
MAL	-0,687	0,917	-0,298	0,553	0,739	1		
CIT	-0,051	0,041	-0,298	0,160	0,271	0,283	1	
SHI	-0,602	0,707	0,031	0,396	0,485	0,644	0,043	1

Observa-se, a título indicativo, com o nível de confiança de 5%, ligações importantes e positivas entre a acidez total e os ácidos málico, shikímico e tartárico. O mesmo ocorre entre os ácidos tartáricos e málico, e entre os ácidos málico e shikímico.

Levando-se em consideração a definição de pH, esta grandeza apresenta correlação negativa com as substâncias citadas. O potássio tem correlação positiva e significativa com os ácidos tartárico e málico.

Geralmente, certas variáveis contribuem com o mesmo tipo de informação. Nestes casos é lógico definir-se um índice único que possa sintetizar o conjunto dos coeficientes de correlação. Isto é possível graças a utilização da Análise em Componentes Principais.

Para eliminar possível distorção da variabilidade devido aos efeitos das unidades das medidas as variáveis são centradas e reduzidas, isto quer dizer que elas são subtraídas de sua média e divididas pelo desvio padrão. $[X = (X - \bar{X} \text{ médio}) / \text{desvio padrão}]$

Esta operação torna-se muito importante no caso deste exemplo para se comparar o potássio com o ácido cítrico, já que estes compostos são expressos em unidades diferentes.

As variáveis de base e o componente principal calculado possuem similaridades que podem ser traduzidas por coeficientes de correlação. O Quadro 6 mostra estes valores.

QUADRO 6 - CORRELAÇÕES ENTRE AS VARIÁVEIS E AS COMPONENTES PRINCIPAIS (COR), E QUALIDADE DE REPRESENTAÇÃO (QUAL)

VARIÁVEIS	AXE 1		AXE 2		AXE 3	
	cor	qual	cor	qual	cor	qual
pH	0,8659	0,749	-0,1188	0,014	-0,3880	0,150
AT	-0,9158	0,838	0,2326	0,054	0,2060	0,042
GP	0,3490	0,121	0,7380	0,545	-0,5360	0,287
K	-0,6749	0,455	-0,3069	0,094	-0,2444	0,059
TAR	-0,7765	0,602	0,0291	0,001	-0,5310	0,282
MAL	-0,9647	0,930	0,0095	0,001	0,0150	0,001
CIT	-0,2447	0,059	-0,6860	0,471	-0,3930	0,150
SHI	-0,7372	0,543	0,3868	0,149	-0,0710	0,005

Observa-se que os coeficientes da primeira componente (CP1) diferem entre si; alguns possuem valores superiores (MAL, AT, pH), outros apresentam valores médios (TAR, K, SHI) e outros mais fracos (GP, CIT).

Neste caso admite-se que eles ponderam as variáveis de base de forma diferente. Os mais importantes participam com maior intensidade na construção do eixo 1 (CP1), enquanto que os mais fracos praticamente não interferem, e as médias sem serem negligenciadas não apresentam grande participação.

Obtém-se então, uma primeira componente principal, que apresenta ela mesma um valor próprio de 4.30. Este valor próprio é igual a inércia da nuvem de pontos da população projetada sobre a componente principal. Por definição, cada variável tem uma variância igual a 1, e a variância de população (soma das variâncias das variáveis) é igual ao número total de variáveis, que no presente exemplo é igual a 8. Neste caso é possível exprimir a percentagem máxima, isto é 53,8% de variância representada pela primeira componente principal.

Logo pode-se dizer que a primeira componente principal, que é uma variável sintética, traduz aproximadamente a metade da variabilidade total do conjunto das oito variáveis de base.

A variabilidade representada pela CP1 é de 53,8%, faltando 46,2% que não foram ainda expressos. Deve-se então procurar uma segunda componente (CP2), que possua propriedades análogas à primeira componente mas que lhe seja ortogonal; isto é, que reagruppe as variáveis capazes de representar a maior parte da variância residual (não definida pela CP1), correlacionadas entre elas e não correlacionadas com as da primeira componente principal.

O número de componentes levadas em consideração para explicar as variações de uma população depende de cada caso preciso.

Segundo ESCOFIER & PAGES (1), deve-se reter na interpretação de uma análise os fatores que se sabe interpretar com clareza. Neste exemplo julgou-se interessante utilizar as três primeiras componentes principais que juntas explicam 82,7% da variabilidade total. As componentes principais são materializadas por eixos na representação gráfica.

O Quadro 7 mostra os valores próprios e a contribuição à variação total das três primeiras componentes principais.

Observa-se (Quadro 6) para a segunda componente principal, que o grau provável (GP) e o ácido cítrico têm as maiores contribuições na sua construção.

Para o eixo 3, é ainda o grau provável e o ácido tartárico que melhor definem esta variabilidade.

QUADRO 7 - VALORES PRÓPRIOS DOS COMPONENTES PRINCIPAIS E
PORCENTAGEM DE VARIAÇÃO EXPLICADA POR CADA EIXO

	CP1	CP2	CP3
valores próprios	4.302	1.329	0.983
%	53.8	16.6	12.3

Depois do estudo das variáveis, inicia-se o estudo das observações, utilizando-se evidentemente os resultados precedentes calculados para as variáveis.

As distâncias entre os indivíduos, em uma ACP, são determinadas pelo cálculo da distância EUCLIDIANA que leva em consideração os diferentes planos do espaço.

As coordenadas dos indivíduos sobre as componentes principais estão contidas no Quadro 8.

Neste quadro, observa-se que para a primeira componente, os 10 primeiros indivíduos apresentam projeções positivas e que para as 16 restantes estes valores são negativos, exceto para os números 15, 17 e 21 que são próximos de zero.

O simples exame das coordenadas dos indivíduos permite ordenar as observações através da primeira componente. Graças a ela pode-se "classificar" os solos de onde provém as amostras, isto porque: os primeiros, provenientes da parcela arenosa (valores positivos) foram obtidos a partir de uvas que apresentavam melhor maturação, estes são menos ácidos que os mostos da parcela argilosa que apresentaram valores negativos.

No que concerne às componentes 2 e 3, a diferenciação dos valores das coordenadas não é tão simples como foi para a primeira. Naturalmente uma representação gráfica vai ser de grande utilidade para se examinar melhor onde se situam as diversas observações.

A Figura 1 mostra o círculo de correlações que é a representação gráfica da extremidade dos setores representativos de cada variável e sua relação com as componentes principais.

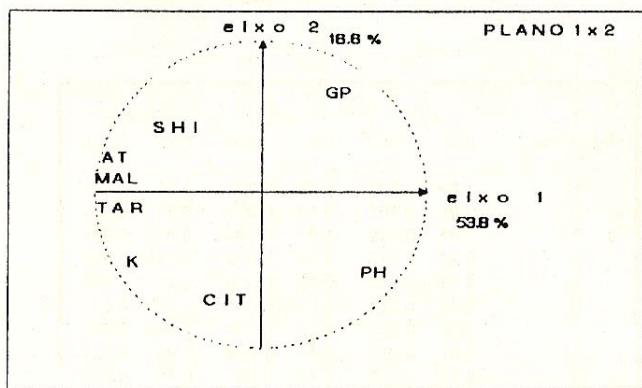
A Figura 2 representa a distribuição dos 26 indivíduos em relação a estas mesmas componentes.

QUADRO 8 - COORDENADAS DOS INDIVÍDUOS SOBRE OS EIXOS PRINCIPAIS E SUA QUALIDADE DE REPRESENTAÇÃO

Indivíduos	AXE 1		AXE 2		AXE 3	
	coor	qual	coor	qual	coor	qual
1	1,59	0,332	0,08	0,001	2,19	0,631
2	0,85	0,161	0,12	0,003	0,61	0,084
3	0,35	0,038	-1,00	0,304	0,46	0,064
4	0,29	0,017	-1,19	0,275	0,71	0,098
5	3,01	0,792	0,82	0,060	0,63	0,034
6	2,69	0,854	0,72	0,061	-0,16	0,003
7	2,76	0,766	1,22	0,150	0,20	0,004
8	2,67	0,919	0,51	0,034	-0,46	0,027
9	4,12	0,895	0,78	0,032	0,49	0,012
10	2,80	0,788	0,24	0,006	-1,08	0,116
11	-1,94	0,408	1,65	0,297	0,09	0,001
12	-3,27	0,830	0,88	0,060	-0,73	0,042
13	-1,98	0,428	1,43	0,221	-0,17	0,003
14	-2,24	0,608	0,63	0,049	-1,06	0,136
15	0,11	0,003	0,33	0,030	-1,65	0,737
16	-0,28	0,170	-0,30	0,020	-1,50	0,490
17	0,15	0,007	-0,43	0,057	-1,68	0,855
18	-0,18	0,008	-0,46	0,052	-1,74	0,742
19	-0,09	0,002	-1,74	0,860	0,33	0,030
20	-0,78	0,087	-2,45	0,855	0,01	0,001
21	0,27	0,011	-1,95	0,554	0,76	0,085
22	-0,63	0,051	-2,48	0,785	0,03	0,001
23	-2,08	0,461	1,28	0,175	1,39	0,205
24	-2,18	0,700	0,94	0,130	0,32	0,015
25	-2,66	0,715	0,43	0,018	1,22	0,151
26	-3,36	0,831	-0,09	0,001	0,83	0,050

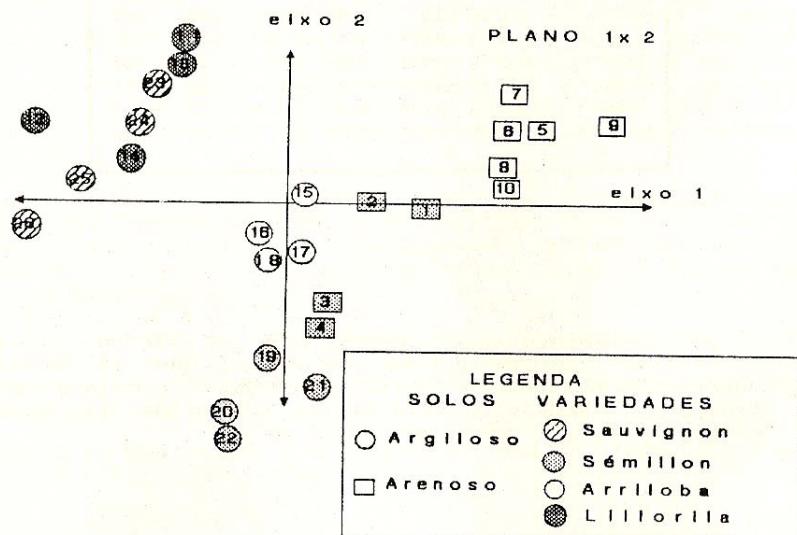
O posicionamento das observações na Figura 2, permite visualizar como se situam os mostos, uns em relação aos outros. Os números de 5 a 10 estão muito próximos, e isto é igualmente verdadeiro para outros grupos de observações: do número 11 ao 14 e do número 23 ao 26.

FIGURA 1 - REPRESENTAÇÃO DAS VARIÁVEIS - CÍRCULO DE CORRELAÇÕES



LEGENDA
 pH = pH AT = acidez total GP = grau provável K = potássio
 TAR = ácido tartárico MAL = ácido málico CIT = ácido cítrico
 SHI = ácido shikímico

FIGURA 2 - REPRESENTAÇÃO DOS INDIVÍDUOS



Entretanto, este gráfico é um plano, logo não se pode afirmar que os pontos vizinhos no plano 1×2 o serão igualmente em outras dimensões.

Para remediar este problema, que pode induzir a erros, deve-se levar em consideração os valores dos coeficientes de correlação ao quadrado, para a representação das variáveis e do cosseno do ângulo formado pelo ponto e sua projeção sobre o eixo ao quadrado, para os indivíduos. Quanto mais próximos ao valor 1, melhor é a representação do ponto no plano. Neste caso, para cada eixo, uma coluna suplementar nos quadros 6 e 8 indicam a qualidade de representação.

No plano 1×2 (Figura 2) os pontos 5 a 10 estão próximos, o que permite dizer que estes pontos têm as mesmas características em relação à primeira componente principal. Efetivamente, neste exemplo, a semelhança destes mostos é devida ao fato de pertencerem ao mesmo tipo de solo e a mesma variedade (arenoso, Arriloba).

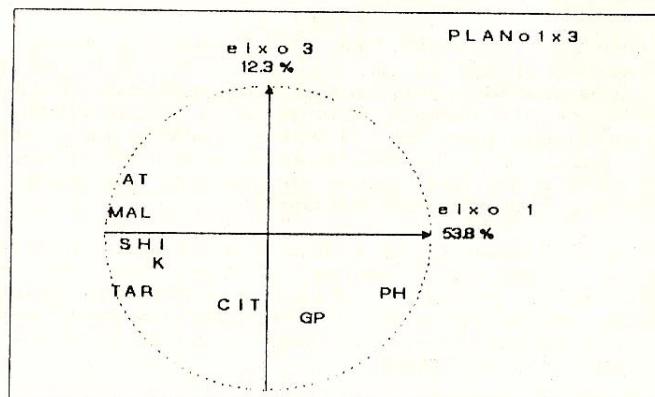
Para os indivíduos de 19 a 22 (Sémillon, Argila), sua discriminação em relação ao grupo Sémillon - arenoso (1 a 4) é posta em evidência no plano 1×2 , no eixo 2, devido aos pequenos valores do grau provável e os fortes teores em ácido cítrico.

Na outra extremidade do eixo 1, no plano 1×2 (Figura 2) constata-se a proximidade de um grupo de "indivíduos" que pertencem a um só solo (argila), mas há duas variedades distintas, Lilioirila de 11 a 14 e Sauvignon de 23 a 26. No plano 1×2 estes indivíduos se confundem; porém, os valores da qualidade de representação mostram que os números 11, 13 e 23 não estão bem representados pelas duas primeiras componentes principais. Isto indica que se deve observar o que se passa nos outros planos (1×3) e, se necessário (2×3), levando-se em consideração que a cada vez a explicação da variabilidade das amostras é menor.

O plano 1×3 (Figuras 3 e 4), permite efetivamente separar melhor os mostos de Sauvignon dos de Lilioirila, exceto o par 11 e 24 que estão muito próximos, mas que ainda apresentam uma qualidade de representação fraca para este plano. Estas duas variedades são separadas por pequenas diferenças em grau provável, ácido cítrico, pH e ácido tartárico.

O grupo de indivíduos do 15 ao 18 (Arriloba, argila) que se encontram no centro do plano 1×2 (Figura 2) e que não se diferenciam bem do grupo do mosto de 1 a 4 (Sémillon, arenoso), está perfeitamente separado pelo eixo 3, no plano 1×3 (Figura 4) em decorrência de seus valores mais elevados em grau provável e ácido tartárico.

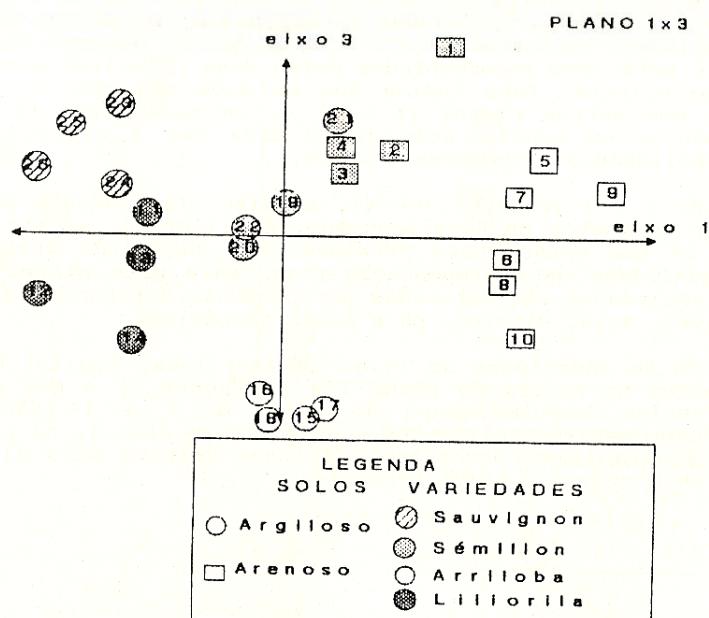
FIGURA 3 - REPRESENTAÇÃO DAS VARIÁVEIS - CÍRCULO DE CORRELAÇÕES



LEGENDA

pH = pH Al = acidez total GP = grau provável K = potássio Mal = ácido málico
 TAR = ácido tartárico CIT = ácido cítrico SHI = ácido shikímico

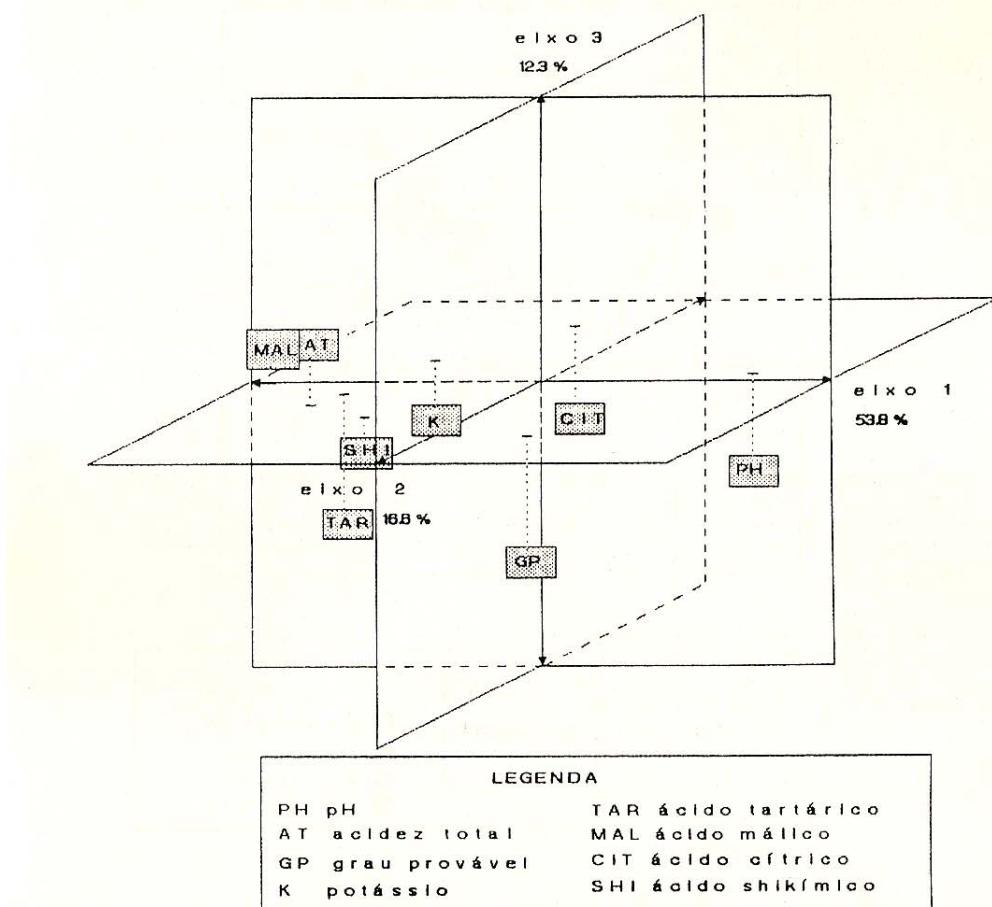
FIGURA 4 - REPRESENTAÇÃO DOS INDIVÍDUOS



A interpretação de dois planos é indispensável para se obter aproveitamento completo da gama de informações que esta ACP pode fornecer.

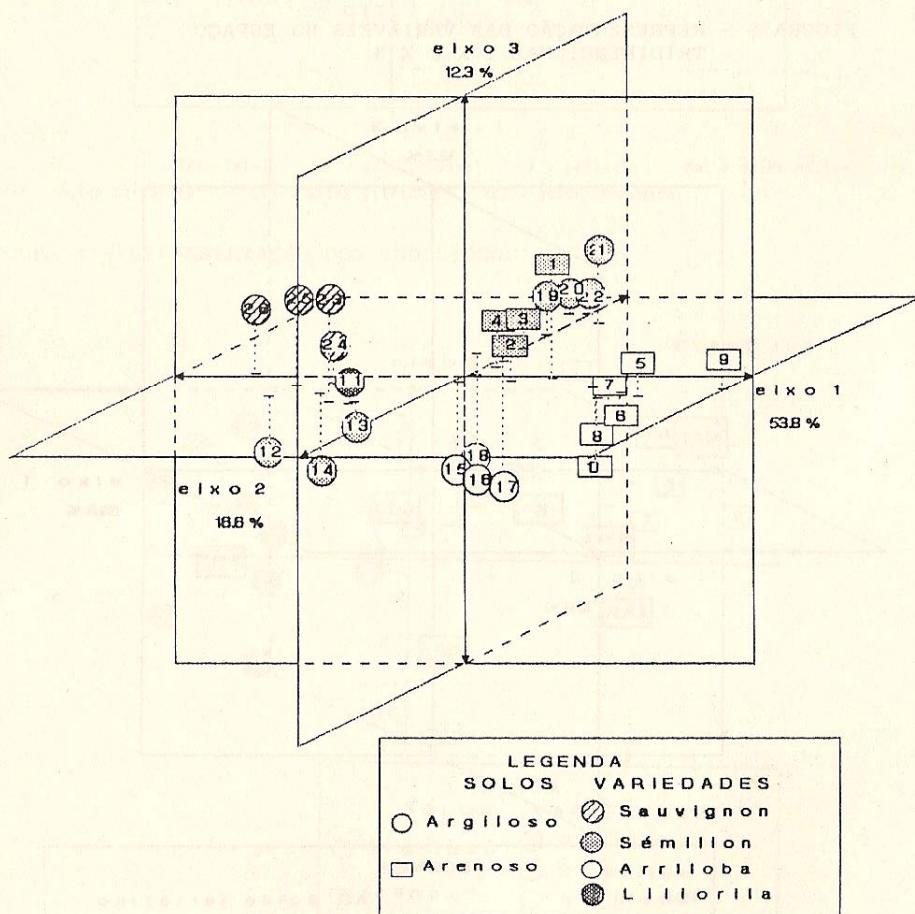
A maneira mais eficaz de auxiliar a compreensão da interpretação em mais de uma dimensão e de reduzir as deformações da representação devidas a projeção sobre o plano, é realizar uma terceira dimensão, levando em conta cada indivíduo, ou variável, suas coordenadas sobre um terceiro eixo, como mostra a técnica desenvolvida por ROSIER (3). Esta maneira de representar torna o gráfico mais claro, permitindo a separação dos pontos das variáveis em espaço tridimensional (Figura 5).

FIGURA 5 - REPRESENTAÇÃO DAS VARIÁVEIS NO ESPAÇO TRIDIMENSIONAL 1 X 2 X 3



A Figura 6 mostra efetivamente que os mostos são perfeitamente separados em função dos solos e das variedades pela sua composição em ácidos orgânicos, acidez total, grau provável e pH, provando que as amostras provenientes da parcela arenosa são menos ácidas e mais ricas em açúcar que as da parcela argilosa.

FIGURA 6 - REPRESENTAÇÃO DOS INDIVÍDUOS - ESPAÇO TRIDIMENSIONAL 1 X 2 X 3



3 CONCLUSÃO

A ACP é um método que permite colocar em evidência, graficamente, as informações essenciais contidas em um conjunto de dados. Seu maior mérito é permitir que se veja mais claro, não para cada variável isolada, mas para um conjunto de variáveis simultaneamente. Por esta razão adapta-se com perfeição ao estudo dos mostos e vinhos, que são compostos de um grande conjunto de substâncias.

ANEXO 1 - AUXÍLIO PARA INTERPRETAÇÃO DAS ANÁLISES MULTIDIMENSIONAIS

Para facilitar a interpretação das análises multidimensionais deve-se levar em consideração as seguintes regras práticas:

- 1 O primeiro passo consiste em verificar a % de variação da população, explicada para cada eixo.
- 2 Estudo das Variáveis
 - 2.1 Para um eixo dado, as variáveis que apresentarem os maiores coeficientes (em valor absoluto) são as que mais contribuem na formação dos eixos.
 - 2.2 Se duas variáveis apresentarem correlação negativa, quando estiverem na mesma direção, mas em oposição sobre o eixo, uma aumentará quando a outra diminuir.
 - 2.3 A proximidade entre duas variáveis, sobre os bordos do círculo de correlação indica forte correlação positiva entre elas, e comportamento semelhante na formação dos eixos.
 - 2.4 A proximidade de uma ou mais variáveis da origem dos eixos, se traduz por pequena participação destas variáveis na construção dos eixos, elas fornecem pouca informação quanto a variação da população em relação a estes eixos.
- 3 Estudo dos Indivíduos
 - 3.1 A proximidade ou o distanciamento dos indivíduos, levando-se em consideração a qualidade de representação, indica suas semelhanças ou diferenças. A posição de um indivíduo no centro do plano indica que ele é considerado médio em relação às informações sobre este eixo.

ABSTRACT

Multivariate statistics analysis, as is the case of the Principal Component Analysis (ACP), fit perfectly for interpretation of enology results. The musts and wines can be defined by the conjunct of its constitutive characteristics. Small differences are not significative in each one of its components and can effectively differentiate each wine from the other. The ACP simplify the comprehension of a complex problem with a minimum loss of informations. This work presents a practical form of an ACP interpretation carried out to evaluate the organic acid composition of grape musts from four different grapevine varieties and from two types of soils.

REFERÊNCIAS BIBLIOGRÁFICAS

- 1 ESCOFIER, B., PAGES, J. Analyses factorielles simples et multiples: Objectifs, méthodes et interprétations. 10 ed. Paris : Dunod, 1990. 266 p.
- 2 PHILIPPEAU, G. Comment interpréter les résultats d'une analyse en composantes principales? Paris : STAT-ITCF Institut Technique des Céréales et des Fourrages, INRA, 1986. 63 p.
- 3 ROSIER, J. P. Interpretation des caractères analytiques et sensoriels de vins blancs de la Région des Graves en fonction de certains facteurs cultureaux de la vigne. Bordeaux, 1992. 260 p. Thèse de Docteur, Université de Bordeaux II.