

ESTUDO ESTATÍSTICO DE VARIÁVEIS DE UMA VINHA NUM SOLO GRANÍTICO.

Ana Silva⁽¹⁾, Ana Braga⁽²⁾, Isabel Araújo⁽³⁾, Teresa Mota⁽⁴⁾, José Oliveira⁽⁵⁾, Jorge Oliveira⁽⁶⁾

⁽¹⁾ Universidade do Minho

Campus de Gualtar, Braga, Portugal

anaritaveirasilva@gmail.com

⁽²⁾ DPS, Universidade do Minho

⁽³⁾ Vinalia – Soluções de Biotecnologia para a Vitivinicultura, Lda.

⁽⁴⁾ Quinta Campos de Lima – Unipessoal Lda.

⁽⁵⁾ IBB – Institute for Biotechnology and Bioengineering, Centre of Biological Engineering

⁽⁶⁾ Sinergeo – Soluções Aplicadas em Geologia, Hidrogeologia e Ambiente, Lda.

ABSTRACT

The Agrocontrol project aims to identify soil variables that influence the development of the vineyard, grape quality and therefore wine quality.

Application of Principal Component Analysis allowed to "reduce" the problem dimensionality in relation to soil variables analyzed and to identify those that most influence the total variability of the data. The Cluster Analysis allowed grouping, in each study plot, the areas with similar characteristics concerning these variables.

The determination of correlations between the main components of soil and the physico-chemical properties of must and vigor and productivity of the vineyard allowed to identify which soil variables most affect the balance of the vine growing and wine quality for one year study.

RESUMO

O Projecto Agrocontrol tem como objectivo identificar as variáveis do solo que influenciam o desenvolvimento da videira, a qualidade da uva e por conseguinte a qualidade do vinho.

A aplicação da Análise de Componentes Principais permitiu "reduzir" a dimensionalidade do problema relativamente às variáveis do solo analisadas e identificar aquelas que mais interferem na variabilidade total dos dados. A Análise de Clusters permitiu agrupar, na parcela, zonas com características semelhantes ao nível destas variáveis.

A determinação de correlações entre as componentes principais do solo e as características físico-químicas do mosto e a produtividade e vigor da videira permitiu identificar quais as variáveis do solo que mais interferem no equilíbrio vegetativo da videira e na qualidade do vinho para um ano de estudo.

INTRODUÇÃO

Os factores ambientais, nomeadamente o clima e o solo intervêm de uma forma inequívoca na qualidade das uvas e consequentemente na do vinho. Estes factores abióticos, além de poderem ser limitantes ou mesmo impeditivos do estabelecimento da cultura, são responsáveis por uma grande diversidade de situações (Clímaco, 1997).

O conhecimento do solo é essencial para melhorar a rentabilização dos sistemas produtivos e minimizar os processos de degradação deste recurso. A informação detalhada sobre os solos de aptidão vitícola deve ser um instrumento de gestão essencial para a rentabilização dos investimentos na produção vitivinícola.

A Estatística Multivariada reúne um conjunto de métodos que permitem a análise simultânea de dados caracterizados por duas ou mais variáveis correlacionadas entre si (Reis, 2001).

Estes métodos têm-se revelado poderosos na manipulação de dados com muitas variáveis nos diversos ramos da actividade científica.

A Análise de Componentes Principais (ACP) permite uma redução do número de variáveis, fazendo com que o fenómeno em estudo seja representado da maneira mais simples possível, sem sacrificar informações valiosas e tornando as interpretações mais simples (Ferreira, 1996). As novas variáveis designam-se por Componentes Principais (CP) e podem ser usadas posteriormente como “índices” ou indicadores que resumem a informação disponível nas variáveis originais (Maroco, 2007).

A Análise de Clusters (AC) tem como objectivo construir grupos ou *clusters* de tal forma que os objectos dentro do mesmo grupo são mais semelhantes do que objectos situados em grupos diferentes (Branco, 2004). Obtêm-se assim grupos homogéneos ou compactos relativamente a uma ou mais características comuns (Maroco, 2007).

O Projecto Agrocontrol co-financiado pelo “ON.2 – O Novo Norte” e QREN através do Fundo Europeu de Desenvolvimento Regional (FEDER), teve como objectivo identificar e estudar múltiplas variáveis que determinam o comportamento físico e químico dos solos, que por sua vez influenciam o desenvolvimento da videira, a qualidade final da uva e por conseguinte a qualidade dos vinhos. Este conhecimento pode ser utilizado para otimizar a produção, de forma a conseguir um produto com características próprias, em consonância com adequadas tecnologias de vinificação.

A parcela em estudo situa-se na Estação Vitivinícola Amândio Galhano (EVAG) na Quinta Campos de Lima, no concelho dos Arcos de Valdevez, e está inserida numa região apta à produção do DOC Vinho Verde. A vinha em estudo compreende apenas cv Vinhão, encontrando-se explorada em Modo de Produção Biológico (MPB), certificada pela EcoCert Portugal.

A partir dos dados obtidos em campo e laboratório, e recorrendo a técnicas de estatística multivariada e determinação de correlações, identificaram-se parâmetros do solo que influenciaram o desenvolvimento da planta em 2010 e a qualidade das uvas e respectivo vinho.

MATERIAIS E MÉTODOS

Na parcela em estudo foram georreferenciados 45 pontos, com espaçamento regular, resultando numa malha de 15 linhas por 3 colunas (Fig.1).

As amostras de solo foram recolhidas em cada um dos 45 pontos georreferenciados à profundidade de 15-20 cm e para cada uma foram determinados os seguintes parâmetros: pH em extracto aquoso (pH); matéria orgânica, em % (MO); densidade aparente, em gramas de solo seco por volume de solo não perturbado (DA); fracção fina, em % (FF); fracção grosseira, em % (FG); fósforo assimilável, em $\mu\text{g/g}$ (P2O5); potássio assimilável, em $\mu\text{g/g}$ (K2O); cálcio assimilável, em $\mu\text{g/g}$ (Ca); magnésio assimilável, em $\mu\text{g/g}$ (Mg); azoto total, em % (AzT); níquel, em $\mu\text{g/g}$ (Ni); crómio, em $\mu\text{g/g}$ (Cr); cádmio, em $\mu\text{g/g}$ (Cd); nitratos, em $\mu\text{g/g}$ (N); boro, em $\mu\text{g/g}$ (B) e capacidade de troca catiónica, em m.e./100g (CTC).

A produção e vigor das videiras foram avaliados em consonância com os 45 pontos georreferenciados. Entre as componentes da produção das videiras foi avaliado: número de cachos por videira (Ncachos); peso médio do cacho por videira, em kg (Pcacho_kg) e peso de uvas por videira, em kg (uvas_kg_vid). O vigor das videiras foi definido através das variáveis: número de varas por videira (Nvaras); peso das varas por videira, em kg (Pvaras_kg) e peso médio da vara por videira, em gramas (Pvara_g).

Em relação ao mosto, foram recolhidas nove amostras compósitas de bagos de acordo com as nove localizações esquematizadas na Fig. 1. Para a caracterização do mosto foram efectuadas as seguintes análises de acordo com o Regulamento (CEE) nº 2676/90, de 17 de Setembro de 1990: pH do mosto (pH mosto); acidez total, em g/dm^3 (Acidez_T); ácido málico, em g/dm^3 (Acido_malico); ácido tartárico, em g/dm^3 (Acido_tartarico); açúcares, em g/dm^3 (Açúcares) e teor de álcool provável, em % vv (TAP).

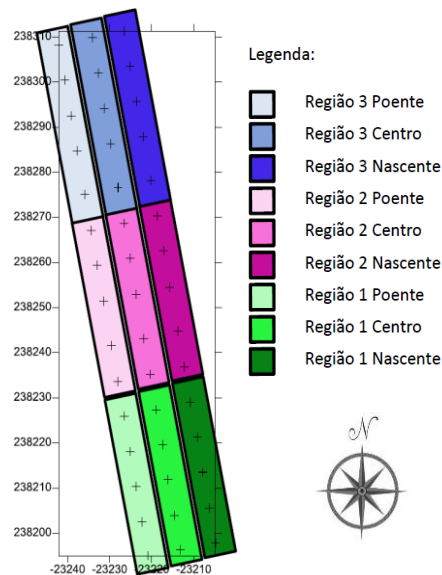


Fig. 1 – Parcela com localização dos 45 pontos georreferenciados

Para a análise estatística dos dados recorreu-se a técnicas de estatística multivariada tais como ACP e AC e à determinação de correlações. As diferenças estatísticas das variáveis de acordo com os clusters foram avaliadas por análise da variância (ANOVA). A normalidade dos dados foi verificada com o teste Kolmogorov-Smirnov com correcção de Lilliefors e a homogeneidade das variâncias pelo teste de Levene, ambos ao nível de significância de 5%. Sempre que estes pressupostos falharam foi aplicado o teste não paramétrico de Kruskal-Wallis.

Os softwares utilizados foram o SPSS Statistics 19 e o Surfer 9.

RESULTADOS E DISCUSSÃO

Em relação às variáveis do solo em estudo verifica-se que há características do solo que apresentam maior variabilidade e que as variáveis MO, FF, FG, P2O5, Mg, AzT, Cd e B apresentam *outliers*. Podemos mesmo observar a existência de *outliers* severos, assinalados com * na Fig. 2, relativamente às variáveis MO, AzT, Cd e B.

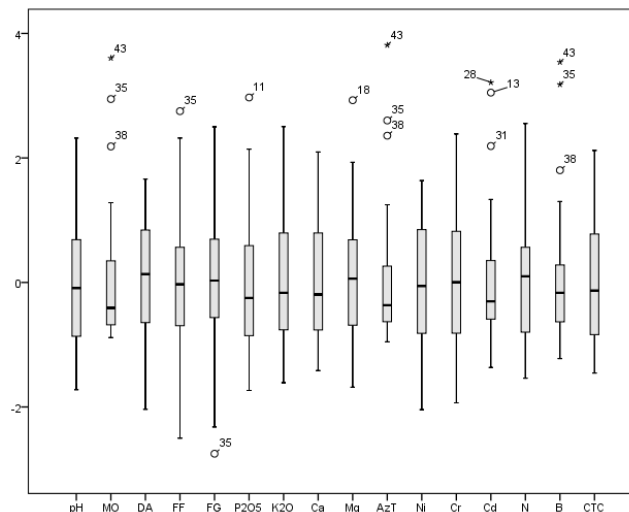


Fig. 2 – Boxplots relativos às variáveis do solo

A matriz de correlações das variáveis do solo e o cálculo do seu determinante (muito próximo de zero) permite concluir que as variáveis estão correlacionadas entre si, pelo que muitas delas poderão inclusivamente estar a dar a mesmo tipo de informação.

Desta forma aplicou-se uma ACP com o objectivo de “reduzir” a dimensionalidade do problema, esperando-se que as primeiras componentes principais expliquem grande parte da variabilidade dos dados iniciais.

Procedeu-se à derivação das componentes principais começando por calcular a proporção de variância explicada pelas novas variáveis e seus *loadings*. Os resultados obtidos encontram-se nas Tab. 1 e 2.

Tab.1 – Variância Total Explicada

Component	Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %
1	4,167	26,042	26,042
2	3,523	22,021	48,063
3	2,650	16,560	64,623
4	1,485	9,283	73,906

Extraction Method: PCA

Tab.2 – Matriz das CPs com rotação

	Component			
	1	2	3	4
pH	0,773	-0,195	-0,284	-0,207
MO	-0,071	0,955	0,234	-0,054
DA	-0,104	-0,015	-0,012	0,810
FF	-0,213	0,293	0,820	0,125
FG	0,213	-0,293	-0,820	-0,125
P2O5	0,612	-0,387	-0,006	0,307
K2O	0,341	-0,008	0,518	0,067
Ca	0,955	0,125	-0,077	-0,067
Mg	0,781	-0,182	0,081	0,246
AzT	0,026	0,961	0,199	-0,086
Ni	-0,198	0,427	0,511	-0,421
Cr	-0,002	-0,098	-0,705	0,163
Cd	0,395	-0,107	0,105	0,497
N	0,557	0,410	-0,133	0,358
B	-0,050	0,948	0,249	-0,011
CTC	0,970	0,091	-0,023	-0,028

Extraction Method: PCA

Rotation Method:

Varimax with Kaiser Normalization.

Analisando a coluna relativa à proporção acumulada na Tab. 1, pelo critério da variância total explicada, devemos reter as 4 primeiras CPs, as quais explicam aproximadamente 73,9 % da variabilidade total dos dados.

Na Tab. 2 podem-se analisar os *loadings* para as 4 componentes principais retidas, ou seja, os coeficientes das combinações lineares de cada uma das componentes principais em função das variáveis originais. Desta análise conclui-se que para a 1ª CP as variáveis que têm mais peso são: pH, P2O5, Ca, Mg, N e CTC, todas com coeficientes positivos. Na 2ª CP as variáveis com maior peso são: MO, AzT e B, com coeficientes positivos. Na 3ª CP temos as variáveis: FF, K2O e Ni com coeficientes positivos e FG e Cr com coeficientes negativos. Finalmente, na 4ª CP tem-se as variáveis DA e Cd com coeficientes positivos.

Na Fig. 3, apresenta-se a distribuição espacial dos *scores* de cada uma destas CPs, ou seja, dos valores que cada um dos 45 pontos toma para cada uma destas novas variáveis.

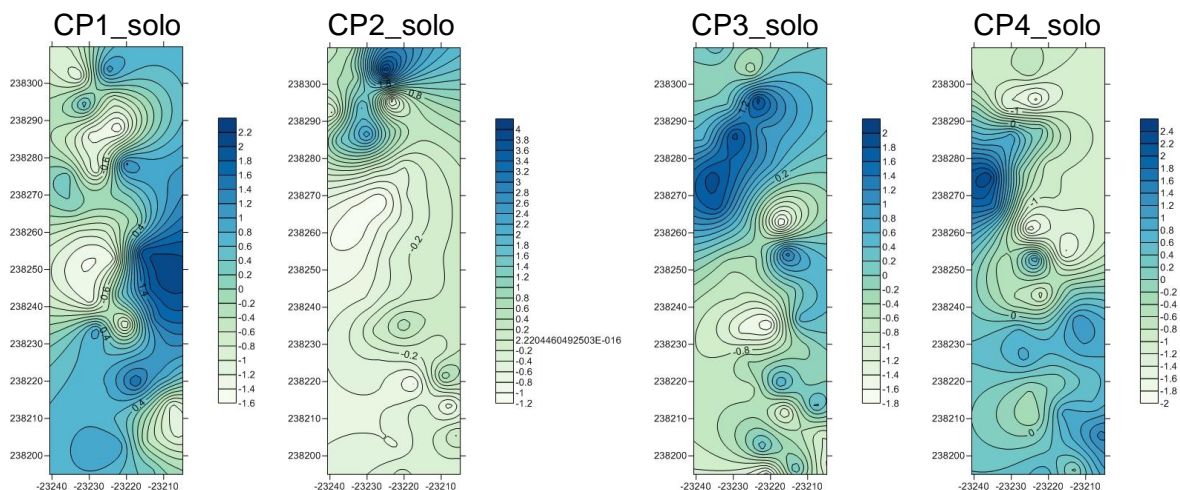


Fig.3 – Distribuição espacial dos scores das componentes principais do solo

Procedeu-se de seguida a uma Análise de *Clusters* com o objectivo de identificar na parcela, zonas com características semelhantes ao nível das variáveis do solo.

A Fig. 4 ilustra o dendrograma obtido de acordo com as novas variáveis independentes obtidas na ACP.

A análise gráfica sugere a divisão da parcela em 4 grupos, os quais se encontram identificados na Fig. 5.

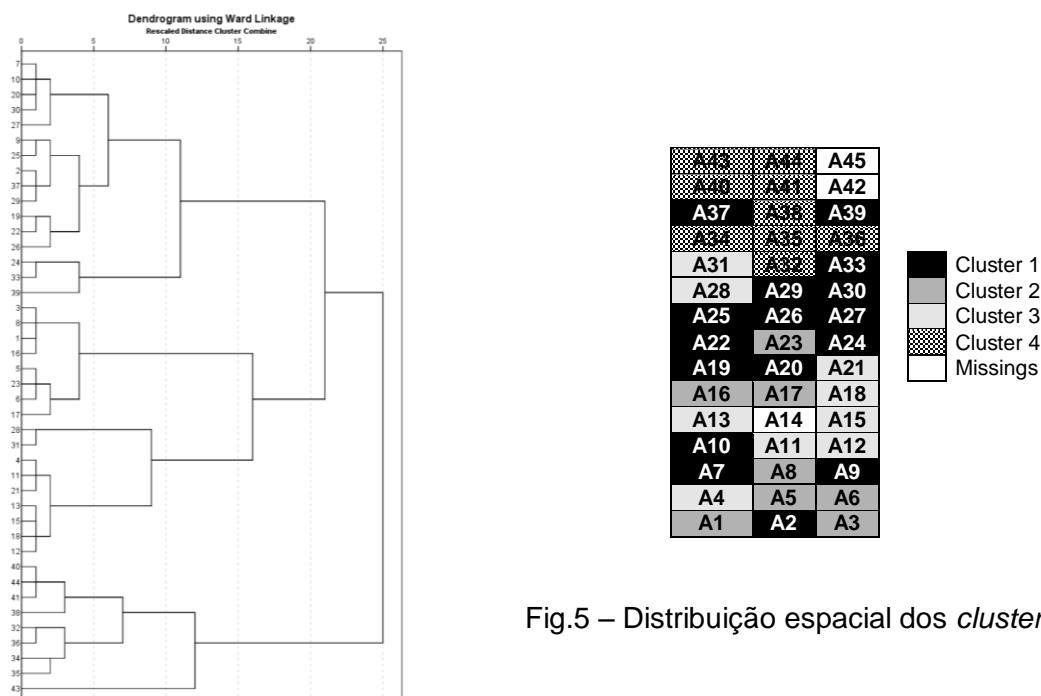


Fig.5 – Distribuição espacial dos clusters

Fig.4 – Dendrograma para as 45 subamostras com base nas componentes principais do solo

Verificou-se que os valores médios de todas as variáveis do solo analisadas apresentam diferenças significativas de *cluster* para *cluster*, à excepção da variável N, como podemos confirmar na Tab. 3.

Tab.3 – Resultados dos testes de comparação para os constituintes do solo

	pH (b)	MO (b)	DA (a)	FF (a)	FG (a)	P2O5 (a)	K2O (a)	Ca (b)	Mg (a)	AzT (b)	Ni (a)	Cr (a)	Cd (b)	N (a)	B (b)	CTC (a)
Cluster 1	5,85	2,42	1,08	67,09	32,91	42,34	82,10	463,13	70,59	0,105	3,77	0,38	0,102	1,21	0,36	10,90
Cluster 2	5,73	2,65	1,21	64,46	35,54	46,52	69,76	426,56	69,75	0,112	2,30	0,66	0,115	1,45	0,38	10,13
Cluster 3	5,95	2,97	1,23	68,34	31,66	57,59	105,47	613,33	86,33	0,130	3,28	0,37	0,154	1,44	0,40	14,20
Cluster 4	5,42	6,22	1,11	73,17	26,83	23,37	90,91	399,17	64,00	0,250	5,26	0,35	0,103	1,34	0,57	9,67
Estatística_F Chi-Square ^(b)	12,971 df=3	21,67 df=3	6,818	4,643	4,643	9,014	3,698	10,482 df=3	8,380	20,07 df=3	14,499	5,212	12,94 df=3	2,422	21,78 df=3	4,574
Valor prova	0,005	0,000	0,001	0,007	0,007	0,000	0,020	0,015	0,000	0,000	0,000	0,004	0,005	0,081	0,000	0,008

^(a) ANOVA

^(b) Kruskal-Wallis

A qualidade do vinho foi avaliada pelas características físico-químicas do mosto anteriormente mencionadas e o equilíbrio vegetativo da videira pelos parâmetros da produtividade e do vigor.

Verificaram-se diferenças significativas nos valores médios das variáveis de *cluster* para *cluster*, apenas para as variáveis ácidos málico e tartárico do mosto e para o peso médio do cacho, o número de varas e respectivo peso, como se pode constatar na Tab. 4.

Tab.4 – Resultados dos testes de comparação para os constituintes do mosto e da videira

	MOSTO						VIDEIRA					
	pHmosto (b)	Acidez_T (a)	Ac_Málico (b)	Ac_Tartárico (b)	Açúcares (b)	TAP (b)	Ncachos (a)	uvas_kg_vid (a)	Pcacho_kg (a)	Nvaras (b)	Pvaras_kg (a)	Pvara_g (a)
Cluster 1	3,29	5,97	1,99	4,61	216,71	12,88	36,50	5,24	0,14	28,86	2,13	73,37
Cluster 2	3,33	5,67	1,75	4,59	215,44	12,80	44,25	4,96	0,11	34,88	2,54	70,57
Cluster 3	3,34	5,86	2,10	4,49	215,54	12,81	44,86	6,45	0,15	38,00	3,55	95,64
Cluster 4	3,34	6,31	2,73	4,25	216,45	12,86	29,88	4,16	0,14	29,63	2,11	71,12
Estatística_F ^(a) Chi-Square ^(b)	2,924 df=3	1,891	17,121 df=3	9,196 df=3	0,730 df=3	0,730 df=3	2,213	1,364	3,166	8,152 df=3	5,266	2,465
Valor prova	0,403	0,150	0,001	0,027	0,866	0,866	0,105	0,271	0,037	0,032	0,004	0,080

^(a) ANOVA

^(b) Kruskal-Wallis

A análise das correlações entre as CPs do solo e as variáveis relativas ao mosto e à videira, apresentada na Tabela 5, identifica como significativas as correlações entre a CP2 do solo e a acidez total e o ácido málico, entre a CP3 do solo e o número de cachos por videira e entre a CP4 do solo e o pH do mosto e o número de varas por videira.

Tab.5 – Correlações significativas entre as CPs do solo e as variáveis do mosto e da videira

	pH_mosto vs CP4_solo	Acidez_T vs CP2_solo	Acido_malico vs CP2_solo	Ncachos vs CP3_solo	Nvaras vs CP4_solo
Pearson C	0,329*	0,350*	0,542**	-0,370*	0,389*
Sig. (2-t)	0,047	0,034	0,001	0,024	0,017

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

A acidez total do mosto e a CP2 do solo que é explicada maioritariamente pelas variáveis MO, AzT e B apresentam uma correlação significativa positiva, ou seja, a valores elevados destes constituintes do solo correspondem maiores valores de acidez total no mosto. O *cluster 4* é o que apresenta valores médios maiores nestas variáveis.

A CP2 do solo também se correlaciona positivamente e de forma significativa com o ácido málico do mosto, as quais também apresentam valores médios mais altos no *cluster* 4.

A CP4 do solo, para a qual têm mais peso as variáveis DA e Cd, está correlacionada positivamente com o pH do mosto. Assim, a valores altos destes constituintes correspondem valores mais elevados no pH do mosto. Os valores médios maiores destas variáveis apresentam-se no *cluster* 3.

A CP4 do solo também se correlaciona positivamente e de forma significativa com o número de varas por videira, as quais apresentam valores médios mais baixos no *cluster* 1.

Finalmente tem-se uma correlação negativa significativa entre a CP3 do solo e o número de cachos por videira. As variáveis FF, K2O e Ni têm *loadings* pesados positivos nesta CP, portanto a valores elevados destes constituintes correspondem poucos cachos por videira, como podemos verificar no *cluster* 4. As variáveis FG e Cr têm *loadings* pesados negativos nesta CP, assim, a valores baixos destes constituintes correspondem poucos cachos por videira, o que também se verifica no *cluster* 4.

CONCLUSÕES

Neste estudo, a qualidade do vinho e o equilíbrio vegetativo da videira foram avaliados através das suas correlações com os parâmetros do solo simplificados em quatro CPs.

Em relação às correlações obtidas entre a CP2 do solo e acidez do mosto, existem já estudos que referem que altos valores de MO no solo induzem alta acidez nos mostos (Delas *et al.*, 1992; Araújo, 2004).

A determinação de correlações entre as características do solo, a qualidade dos mostos e o equilíbrio vegetativo da videira irá permitir, sempre que se justifique, correções diferenciadas para cada local da parcela, efectuando-se deste modo, uma agricultura de uma forma sustentada.

BIBLIOGRAFIA

Araújo I., 2004. Características Aromáticas e Cromáticas Castas Amaral e Vinhão. *Tese de Mestrado*. Universidade do Porto/Universidade Técnica de Lisboa.

Branco J., 2004. *Uma Introdução à Análise de Clusters*. 1ª Edição. Lisboa: Sociedade Portuguesa de Estatística.

Clímaco P., 1997. Influência da cultivar e da maturação da uva e na produtividade da videira (*Vitis vinifera* L.). *Tese de Doutoramento*. Instituto Superior de Agronomia, Universidade Técnica de Lisboa.

Delas J., Molot C., Soyer J.P., 1992. Influence du porte-greffe et de la fertilization azotée sur la composition des baies de Merlot. *IV Simposio Internazionale di Fisiologia della Vite*, Torino, Itália, n: 397-400.

Ferreira D., 1996. *Análise Multivariada*. 1ª Edição. Lavras: Universidade Federal de Lavras, Departamento de Ciências Exactas.

Maroco J., 2007. *Análise Estatística com utilização do SPSS*. 3ª Edição. Lisboa: Edições Sílabo.

Pestana M., 2008. *Análise de Dados para Ciências Sociais – A Complementaridade do SPSS*. 5ª Edição. Lisboa: Edições Sílabo.

Reis E., 2001. *Estatística Multivariada Aplicada*. 2ª Edição. Lisboa: Edições Sílabo.