

Improving early-season estimates of olive production using airborne pollen multi-sampling sites

Helena Ribeiro · Mário Cunha · Ilda Abreu

Received: 13 October 2006 / Accepted: 18 January 2007
© Springer Science+Business Media B.V. 2007

Abstract The aim of this study was to improve the accuracy of aeropalynological models to forecast yields in areas with heterogeneous characteristics by applying principal component analysis to integrate the airborne pollen sampled from more than one trap. The sampling was performed during the past seven years (1998–2004) in the main northeast olive regions of Portugal. Annual crop production was forecasted on the basis of airborne pollen concentration measured at flowering, comparing the performance of three different independent variables: total airborne pollen concentration sampled in each trap and a derived variable that was determined by principal component analysis of the total airborne pollen concentration sampled. The best predictive results

were obtained using a logarithmic relationship with airborne pollen concentration principal component scores describing about 97% of olive fruit production variability over the last seven years. The use of this technique improved the ability of pollen to explain the production interannual variations by about 13%. The comparison between actual reported and the adjusted production showed an average spread deviation of 5%.

Keywords Airborne pollen · Forecasting yield · Multisampling sites · Principal component analysis

Abbreviations

APC Airborne pollen concentration
FOC Vila Nova de Foz-Côa
MPS Main pollination season
PCA Principal components analysis
VAD Valença do Douro

H. Ribeiro · M. Cunha
Secção Autónoma de Ciências Agrárias, Faculdade de
Ciências da Universidade do Porto, Porto, Portugal

H. Ribeiro (✉) · I. Abreu
Departamento de Botânica, Faculdade de Ciências da
Universidade do Porto, Rua do Campo Alegre, 1191,
4150 Porto, Portugal
e-mail: helenaribeiro@fc.up.pt

M. Cunha
CECA-ICETA, Universidade do Porto, Vairao,
Portugal

I. Abreu
IBMC, Instituto de Biologia Molecular e Celular,
Universidade do Porto, Porto, Portugal

1 Introduction

Olive growing is characteristic of the Mediterranean region, playing a social and economic role of paramount importance. This culture is characterized by large production of pollen, great fruit abscission and by a long post-flowering period.

The expansion of this culture leads to the improvement and development of crop management techniques that enhance crop productivity (new growing techniques, pest management, oil transformation efficiency, stock management, marketing strategies, and optimization of human resources).

The forecast of olive yield is a valuable instrument, helping decision making for technical, economic and social operations.

The existence of a relationship between the amount of airborne pollen concentration (APC) and the final yield has been mainly used for crop forecasting in vine and olives (Cour & Campo, 1980; Cunha, Abreu, Pinto, & Castro, 2003; Galán, Vázquez, García-Mozo, & Domínguez, 2004; González-Minero, Candau, Morales, & Tomas, 1998; Msallen, Cour, & Richard, 1996). This procedure enables the integration of all pre-flowering conditions carry-over from the past year(s) and also the easy update of the regional olive surface development, since production areas are often superior to the updating of the regional inventories. The use of aeropalynological data for crop forecasting is based on the principle that the number of flowers is higher in the most productive years; therefore, more pollen is produced, with a higher annual APC as well.

The forecasting models developed over the years on a regional scale use the APC of only one sampling station. However, the area of influence of an aerobiological sampler and the plant's productivity can be affected by the dominant winds in the first case, by soil and genetics in the second case, and by topography, temperature, and precipitation for both cases. Therefore the simultaneous use of more than one sampler per each region can be a good methodology in heterogeneous production areas, improving the accuracy of the aeropalynological models to forecast yields. Also, the post-flowering conditions, when they are not constant in each year, can have a large influence on the crop yield.

The aim of our study was to improve the accuracy of the aeropalynological models for the forecast of olive yields, comparing the performance of a single sampler to a combination of two samplers.

2 Material and methods

2.1 Study area

Olive trees (*Olea europaea* L.) are one of the most important crops in the Mediterranean area, occupying a significant proportion of land use in Portugal, comprising 10% (325000 ha) of the utilised agricultural area (UAA) (INE, 1999).

This research was performed in the north-east of Portugal. This region presents some relief features with a marked continental influence that contributes to very high temperatures in the summer and very low temperatures in winter. According to the reference meteorological series 1961/90, the average minimum temperature in winter varies between 0°C and 2°C, while in summer the average maximum temperature is between 26°C and 32°C. The average annual number of frost days (minimum temperature lower than 0°C) ranged between 70 and 110 while the number of hot days (maximum temperature higher than 35°C) ranged between 5 and 15 days. The mean annual accumulated precipitation varies between 900 mm and 1400 mm, mainly centred in winter (Miranda, Coelho, Tomé, & Valente, 2001).

For this study a division in two regions of production, Douro and Alto Trás-os-Montes, was performed.

In the north-east of Portugal the olive cultivation in 2004 represented 27% and 29% of Portugal's oil and fruit production, respectively. The most common olive varieties are Galega, Carrasquenha, Cordovil, Cobrançosa, and Verdeal.

2.2 Pollen sampling and olive production

Olive airborne pollen concentration (APC) was sampled over a seven-year period (1998–2004) using Cour traps (Cour, 1974) installed in the regions of Valença do Douro (VAD: 41° 08'N, 7° 33'W) and Vila Nova de Foz-Côa (FOC: 41° 0'N, 7° 26'W) at approximately 10 m above ground level. The estimate range of the trap is around 50 km depending on the region topography and climate (Cambon, 1981) (Fig. 1).



Fig. 1 Map of Portugal with the location of the samplers. A – Valença do Douro; B – Vila Nova de Foz-Côa

Pollen grains were trapped on vertical gauze filters impregnated with silicone oil, with a surface area of 400 cm² fixed vertically on a wind vane, which continually orientates the filters according to the wind direction.

After exposure the pollen grains were removed using several chemical treatments, including the destruction of the gauze, the acetolysis technique, which eliminate most of the residual organic material and the cytoplasmatic content of the pollen grains, and the addition of glycerol (Cour,

1974). Olive pollen qualification and quantification was carried out, independently of the pollen concentration, with three regular traverse rows of an optical microscope (X630).

During the flowering period, the filters were exposed for three or four days and the average daily APC was expressed as the number of pollen grains transported per m³ of air. This is possible taking into account the air velocity during each filter exposed period measured with an Enerco 402 anemometer (Cimel Electronics, Paris), located next to the trap. An average coefficient of 20% was used to estimate the amount of pollen since filter efficiency varies from 17 to 24% according to the wind speed.

The main pollination season (MPS) was characterized by the period where the majority of pollen is sampled, without taking into consideration pollen recirculation, and was utilized to determine the total regional pollen index (the sum of the daily APCs) for each studied region. This regional pollen index was used as the independent variable in the statistical analysis. Therefore, the beginning of the main pollination season (MPS) was defined, as Julian days, when the accumulated sum of pollen reached 5% of the total annual amount and the end of the season was the date when the accumulated sum reached 95% of the total annual pollen.

The information of the annual olive fruit production expressed in tons (t) was provided by the Statistical National Institute of Portugal (INE, 2005).

2.3 Statistical analysis

For both regions the type of relationship between observed olive fruit, as a dependent variable, and the total pollen index, as an independent variable, was studied using linear and nonlinear models. Three different independent variables were tested: total APC from Foz-Côa, from Valença do Douro and a derived variable that was obtained using a principal component analysis (PCA). Categorical principal component analysis simultaneously quantifies categorical variables and reduces the dimensionality of the data. The principal goal is to reduce an original set of variables into a smaller set of uncorrelated

components that represent most of the information found in the original variables. This technique assumes linear relationships between numeric variables, as is the case for our pollen data ($r = 0.93$; $P < 0.001$), and was used to estimate a new variable that integrates the correlated total APC sampled in both Vila Nova de Foz-Côa and Valença do Douro sampling sites into a linear combination. The derived variable was obtained as followed: $yj = (\gamma j_1 x_1 + \gamma j_2 x_2)$, where yj are the scores of the principal component, $\gamma j_1 \gamma j_2$ are the variable loadings and $x_1 x_2$ are the total APC from VAD and FOC respectively. All the calculations were performed using the SPSS 14.0 software.

3 Results

During the studied period, *Olea* pollination season in the north-east of Portugal occurred between the middle of May and the end of July, with the earliest start observed in 1998 in both studied regions and the latest in 2000 (Table 1). The earliest end date of the main pollination period was registered in 2003 and the latest in 2000. The pollination season lasted on average 28 days, with a minimal duration of 24 days and a maximal duration of 38 days.

Regarding the total airborne pollen concentration (APC), an interannual variation was observed with the years 1999 and 2001 registering the highest pollen index values for Vila Nova de

Foz-Côa and 1999 and 2004 for Valença do Douro, while the lowest values were measured in 2000 (Fig. 2).

Concerning fruit production the same tendency of interannual variability was observed. The maximum fruit production was registered in 1999 and the lowest in 2000. Yields in 1998–2000 presented a stronger yearly variability while the period 2001–2004 presented lower interannual variation with a four-year average yield of 50 thousand tons for the Douro region and 32 thousand tons for the Alto Trás-os-Montes region (Fig. 3).

A linear model of the type: $\text{Yield} = m * \text{Pollen} + b$ and a logarithmic model of the type: $\text{Yield} = a * \text{Ln}(\text{Pollen}) + c$ best described the relationship between pollen and olive fruit yield.

The results of the principal component analysis (PCA) of the total airborne pollen concentration sampled in both Vila Nova de Foz-Côa and Valença do Douro sampling sites are presented in Table 1. Since some of the score values obtained are negative (minimum -1.56), it was necessary to add a value of 2 to each observation to transform them into positive numbers to allow the establishment of a logarithmic relationship between pollen and yield.

Our results showed that airborne pollen concentration is a good predictor of the total olive fruit production in the north-east of Portugal, explaining 67–97% of this region's production (Table 2). The use of principal component analysis for the Valença do Douro and Vila Nova de

Table 1 Results of the characteristics of the main pollen season (MPS) and of the derived variable obtained using principal component analysis (scores) of the total airborne

pollen concentration sampled at the Vila Nova de Foz-Côa (FOC) and Valença do Douro (VAD) sampling sites

Main pollination season characteristics							
Year	Begin date ^a		End date ^a		Duration		Pollen/m ³ air Scores
	FOC	VAD	FOC	VAD	FOC	VAD	
1998	127	127	165	162	38	35	0.96
1999	140	140	168	164	28	24	3.31
2000	146	143	170	167	24	24	0.44
2001	137	133	161	160	24	27	3.04
2002	134	136	168	167	34	31	1.74
2003	135	132	159	159	24	27	1.74
2004	142	139	166	163	24	24	2.78

^a The dates are in Julian days

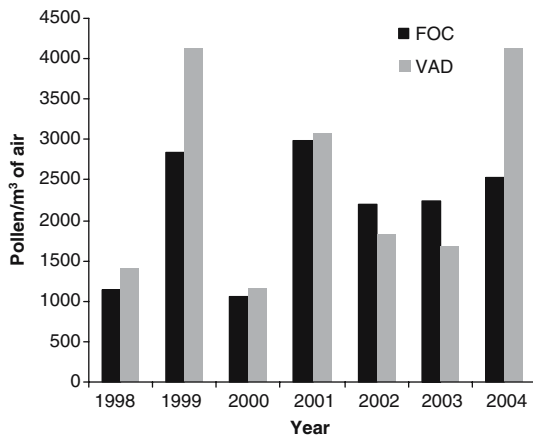


Fig. 2 Total airborne pollen production in the two sampling regions (FOC: Vila Nova de Foz-Côa, VAD: Valença do Douro), expressed as a pollen index, observed from 1998 to 2004

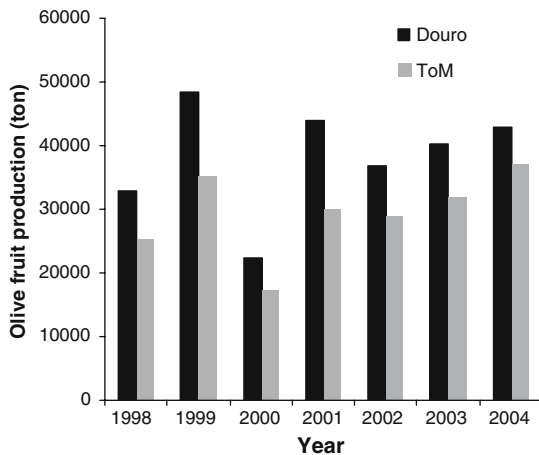


Fig. 3 Total olive fruit production, observed from 1998 to 2004, in the two regions (Douro region and ToM: Alto Trás-os-Montes region) of the north-east of Portugal

Foz-Côa sampling results explain 93% of the total airborne pollen concentration variability.

The highest coefficient of determination (R^2) values were achieved when fruit production (t/ha) is used as the independent variable (between 67% when the total airborne pollen concentration of Valença do Douro is the variable and 97% when the principal component scores are used).

The logarithmic model presented always fits better than the linear one, improving the R^2 values (average of 3%, 9% and 10% for the fruit

production forecast when the total airborne pollen concentration from Vila Nova de Foz-Côa, Valença do Douro and total airborne pollen concentration linear combination of both sites are respectively used). The difference between the observed and predicted fruit production was also lower using the logarithmic model (Table 3).

The use of the principal component scores enhanced the explanation power of the fruit forecasting model compared to the separate use of the total airborne pollen concentration of each sampling point. Regarding fruit production, 97% of Douro and 86% of Alto Trás-os-Montes regions interannual variability can be explained using a logarithmic relationship with the sampled pollen while using a linear relationship only 90% for Douro and 73% for Alto Trás-os-Montes regions can be explained (Table 2).

The smaller differences between the observed and predicted production were always obtained using the logarithmic model and the principal component scores as independent variable. The best fit was achieved for the total fruit production in the Douro region (maximum: 5%, average: 3%). For the Alto Trás-os-Montes region the maximum difference observed was slightly higher (maximum: 15%, average: 6%) (Table 3).

4 Discussion

The pollination season in the studied region occurred later compared to the southern regions of Portugal and earlier compared to the western ones. In fact, according to (Ribeiro, Cunha, & Abreu, 2005), an advance of olive main pollination season beginning and ending dates from the north to the south and from the west to the east of Portugal was observed and is mainly related to geographical and meteorological conditions. The south and west regions have higher average mean temperatures during bud development, in the first case due to a Mediterranean influence, and in the second case due to the proximity of the Atlantic Ocean.

The coefficients of determination (R^2) obtained in our study, with a linear model and only when pollen from one sampler was used, are

Table 2 Results of the linear and nonlinear regression performed between pollen and olive fruit production

Olive fruit production							
Pollen sampler		Linear (Yield = $m \times \text{Pollen} + b$)			Logarithmic (Yield = $a \times \ln(\text{Pollen}) + c$)		
		R^2	m	b	R^2	a	c
Douro ^a	FOC	0.85	10.28***	16200.57**	0.85	18641.51***	-103451.8***
	VAD	0.67	5.52*	24527.68***	0.77	14489.31977***	-73363.68***
	Scores	0.90	7525.96***	23179.93***	0.97	11622.98***	32297.16***
ToM ^a	FOC	0.67	7.07*	14277.90*	0.73	13271.92***	-71451.83***
	VAD	0.67	4.26*	18852.23***	0.74	10949.82***	-54914.05***
	Scores	0.73	5247.60**	18925.38***	0.86	8441.37***	25110.41***

^a The independent variables are FOC: total pollen sampled in Vila Nova de Foz-Côa; VAD: total pollen sampled in Valença do Douro; scores: variable derivate as a linear combination of the total pollen sampled in both FOC and VAD. R^2 – determination coefficient; ToM: Alto Trás-os-Montes region. * – $P < 0.05$; ** – $P < 0.01$; *** – $P < 0.001$. The a and c values are always significant because the confident intervals never contain the value 0

Table 3 Differences in percentage observed between observed and predicted fruit production

		Linear			Logarithmic		
		FOC	VAD	Scores	FOC	VAD	Scores
Olive fruit production—differences between observed and predicted (%)							
Douro	1998	14.8	1.7	7.5	15.2	3.5	3.1
	1999	6.0	2.1	0.7	7.3	2.2	4.5
	2000	-20.6	-38.0	-18.4	-17.1	-28.6	-1.7
	2001	-6.7	5.7	-4.7	-3.9	2.3	-2.8
	2002	-5.1	6.6	1.9	-8.3	4.4	-4.8
	2003	2.3	15.9	9.7	-0.5	14.8	3.6
	2004	1.7	-10.2	-2.8	0.7	-10.1	-3.0
	Average	8	11	7	8	9	3
ToM	1998	11.5	1.7	5.3	12.8	3.2	2.1
	1999	2.6	-3.3	-2.8	3.3	-2.8	0.2
	2000	-25.0	-36.9	-22.4	-20.2	-28.4	-4.8
	2001	-17.9	-6.3	-16.1	-15.7	-9.9	-14.8
	2002	-3.3	8.1	2.9	-6.2	5.7	-3.1
	2003	5.6	18.5	12.1	3.1	17.3	6.7
	2004	13.6	2.1	9.9	12.5	2.6	9.3
	Average	11	11	10	11	10	6

FOC: Vila Nova de Foz-Côa; VAD: Valença do Douro; scores: variable derivate as a linear combination of the total pollen sampled in both FOC and VAD; Douro: Douro region; ToM: Alto Trás-os-Montes region

similar to those obtained in Spain, namely Cordoba ($R^2 = 0.75$) (Galán et al., 2004) and Seville ($R^2 = 0.62$) (González-Minero et al., 1998), in Italy, namely Florence ($R^2 = 0.84$), Prato ($R^2 = 0.69$) (Moriondo, Orlandini, Nuntiis, & Mandrioli, 2001) and Perugia ($R^2 = 0.54$) (Orlandi, Romano, & Fornaciari, 2005) or even in France ($R^2 = 0.74$) (Abid, 1984). However when a logarithmic relationship was used the R^2 values increased, especially when the principal component scores of the airborne pollen concentration (APC) from the two sampling sites were

used as independent variables, increasing by 12% the explanation for the annual olive fruit production. These results may indicate that the precision of pollen as a representative index of future production can be greatly enhanced using a linear combination between two samplers located in different geographic sites. In fact, this technique allowed smoothing of the production outlier in 2000. This year was characterized by low pollen sampling, reflected in fruit production. The differences between declared and adjusted regional production were always higher than 20% in all

cases except when the principal component scores of the two regions total airborne pollen concentration were used in a logarithmic relationship, which was decreased by 3%. Usually a production area is characterized by the existence of a different cultivar with plots located at diverse altitudes and by different local meteorological conditions that affect flower and pollen production and therefore fruit yield. So, for a large production area the use of just one sampler cannot be representative and gives a good estimate of the regional yield production. The use of the total airborne pollen concentration method allowed integration of the pollen sampled in two regions with different local characteristics, improving the yield forecasting at the regional level.

The fit improvement using the logarithmic model compared with the linear model can be explained by an important feature of anemophilous plants, which is the production of a larger amount of pollen than is necessary to pollinate those flowers that are matured into fruit, resulting in fruit-to-flower ratios of less than unity.

Usually the ability of pollen to explain the interannual variability of olive production is still far away from 100%, leaving some uncertainty that is usually explained by other factors (Galán et al., 2004; González-Minero et al., 1998; Moriondo et al., 2001) namely post-flowering factors. Under the Mediterranean conditions this crop is characterized by a long post-flowering period that, together with its extensive management nature in Portugal, influences fruit growing, the maturing period and yields. Consequently this situation enhances the strong interannual variability observed in production. However, it is well known that the flowering stage and pollen production have a crucial function in the determination of crop production, reflecting all past conditions (soil, climate, genetics, diseases and geographical location) affecting the plant from the previous year(s) to this season. These diverse features play a major role in plant productivity. According to several studies (Galán et al., 2004; González-Minero et al., 1998) and to the results herein, the airborne pollen index is the variable with the greatest explanatory capacity for the annual fluctuations in regional olive production. Therefore, in this

study only the airborne pollen was used as the variable and the post-flowering factors were not considered. The main goal was to focus only on the improvement of the variable airborne pollen for the explanation of interannual variability of olive production. This was done by combining the airborne pollen sampled in two different areas, applying principal component analysis to integrate the sampling performed in areas with heterogeneous characteristics. This will allow improvement of the aeropalynological models to forecast yields in a situation where the same post-flowering factors are considered.

The combination of two or more samplers can be useful in situations of scattered olive-growing areas or in high-density sites where differential soil, topography or climate, and especially different dominant wind regimes that affect olive productivity homogeneity and consequently production estimates. For example, in the areas studied that exist two different dominant wind regimes during the olive flowering period. This situation does not allow the maximization of regional pollen sampling because the production areas considered are situated around the sampler in Valença do Douro region and north of the sampler in the Vila Nova de Foz-Côa region. In the former region the wind comes mainly from the south-west to north-west directions while in the latter it is from the north-west to north-east directions. Therefore, these two samplers together allow a sampling of almost 180°, which corresponds to the olive distribution area of north-east of Portugal.

5 Conclusions

Our study demonstrated that airborne pollen concentration during the olive flowering period can be a valuable variable for predicting olive crop production in both studied regions.

The use of a linear combination between two samplers compared with the use of only one sampler represents a very clear improvement in the prediction accuracy of the aeropalynological predictive models for regional olive production.

This methodology increases the potential usefulness of airborne pollen concentration as one of

the independent variables for the forecast of olive production in Portugal and also for other countries.

The accuracy and the early-season indication of crop yields are highly relevant for every operator connected with the olive business (for both economic and technical reasons) and airborne pollen monitoring can give this information, with an average error of 5%, seven months prior to harvest. On the other hand, our results leave some room to improve the explanatory ability of the forecast model by adding post-flowering variables.

Acknowledgments The first author would like to thank the Fundação para a Ciência e Tecnologia SFRH/BD/13148/2003 for support.

References

- Abid, A. (1984). *Contribution à l'étude de la pollinisation de l'olivier (Olea europaea)*. Montpellier: Université de Montpellier II.
- Cambon, G. (1981). Relations entre le contenu pollinique de l'atmosphère et le couvert végétal en Méditerranée occidentale, à Montpellier (France), Valencia (Espagne) et Oran (Algérie). Univ. Sci. Tech., Languedoc, Montpellier.
- Cour, P. (1974). Nouvelles techniques de détection des flux et des retombées polliniques: étude de la sédimentation des pollens et des spores à la surface du sol. *Pollen et Spores*, XVI, 103–141.
- Cour, P., & Campo, M.V. (1980). Prévision de récoltes à partir de l'analyse du contenu pollinique de l'atmosphère. *Comptes Rendus de l'Académie des Sciences série D*, 290, 1043–1046.
- Cunha, M., Abreu, I., Pinto, P., & Castro, R. (2003). Airborne pollen samples for early-season estimates of wine production in a mediterranean climate of Northern Portugal. *American Journal of Enology and Viticulture*, 54, 189–194.
- Galán, C., Vázquez, L., García-Mozo, H., & Domínguez, E. (2004). Forecasting olive (*Olea europaea*) crop yield based on pollen emission. *Field Crops Research*, 86, 43–51.
- González-Minero, F., Candau, P., Morales, J., & Tomas, C. (1998). Forecasting olive crop production based on ten consecutive years of monitoring airborne pollen in Andalusia (southern Spain). *Agriculture, Ecosystems and Environment*, 69, 201–215.
- INE. (1999). Recenseamento Geral à Agricultura.
- INE. (2005). Produção Vegetal das principais culturas (1998–2004), In I. N. Estatística, (Ed.).
- Miranda, P., Coelho, F. E. S., Tomé, A. R., & Valente, M. A. (2001). 20th Century portuguese climate and climate scenarios. In F. D. Santos, et al., (Eds.), *Climate change in Portugal. Scenarios, impacts and adaptation measures - SIAM. Executive summary and conclusions*. (pp. 23–84). 1st edition ed. Lisbon: Gradiva.
- Moriondo, M., Orlandini, S., Nunttiis, P. D., & Mandrioli, P. (2001). Effect of agrometeorological parameters on the phenology of pollen emission and production of olive trees (*Olea europaea* L.). *Aerobiologia*, 7, 225–232.
- Msallen, M., Cour, P., & Richard, P. (1996). Prévision des productions oléicoles par la méthode aéropalynologique. étude du flux pollinique d'olivier en Tunisie et dans quelques pays du nord de la Méditerranée. *Olivae*, 61, 53–56.
- Orlandi, F., Romano, B., & Fornaciari, M. (2005). Relationship between pollen emission and fruit production in olive (*Olea europaea*). *Grana*, 44, 98–103.
- Ribeiro, H., Cunha, M., & Abreu, I. (2005). Airborne pollen of *Olea* in five regions of Portugal. *Annals of Agriculture and Environmental Medicine*, 12, 317–320.