Aerobiologia **18:** 23–31, 2002.

An automated system for surveying and forecasting *Olea* pollen dispersion

Pablo J. Hidalgo[1], Antoine Mangin[2], Carmen Galán[1], Odile Hembise[2], Luis M. Vázquez[1] & Oscar Sanchez[2]
[1]*Department of Plant Biology, University of Córdoba, Campus Universitario de Rabanales, 14071-Córdoba, Spain (E-mail: bv2hifep@uco.es);* [2]*ACRI-ST, 260 Route du Pin Montard, 06904-Valbonne, France (E-mail: am@acri.fr)*

Received 5 February 2001; accepted in final form 16 October 2001

Key words: automated system, forecast, neural networks, *Olea europaea*, pollen emission, phenology, thermal sum

Abstract

An automated system for *Olea* pollen surveying and forecasting is being developed for the province of Cordoba, Spain, within the framework of the European ASTHMA project. Required input data have been split into three categories: static data, calibration parameters and dynamic data. A 200 × 200 km potential pollen production map, centred on the city of Cordoba, was drawn up using digital high-resolution satellite data and reflecting average plant density, as derived from the observed spatial separation of trees in 28 different locations. Based on archives of pollen data from 1982, the start of the pollen season was determined by applying both heat accumulation methods and neural network analysis including further meteorological parameters. The forecast and survey system is based on a set of two elementary atmospheric physics modules and biological models: emission and dispersion modules. The system can run in different modes, allowing forecasting of the flowering start-date and, from that date, charting the emission and dispersion of pollen concentrations in the area. Also, wind and temperature fields are regularly computed and updated using deterministic numerical models to chart the spatial and temporal concentration of *Olea* pollen.

Introduction

The olive (*Olea europaea* L.) is a long-lived tree cultivated for its fruits and oil. The surface area devoted to olive production in Cordoba province accounts for almost 22% of the total surface area, increasing gradually towards the south. The overall olive production area in 1997 was 313,971 ha. (Andalusia Statistical Yearbook, IEA). The city of Cordoba is located in the southwest of the Iberian Peninsula (Figure 1). The city is 120 m above sea level, in the valley of the river Guadalquivir; it has a Mediterranean climate, with a touch of continentality. Winters are short and mild and summers are long and hot. The average annual temperature is 18 °C and annual rainfall is close to 600 mm. The large area given over to olive-growing leads to a high concentration of olive pollen in the atmosphere. Sensitivity to olive pollen grains is among the most frequent allergenic reactions for the local population (Dominguez el al., 1993). With regards to the risk threshold, Florido et al. (1999) have reported that a concentration of around 400 pollen grains/m^3 is sufficient to provoke symptoms in all patients clinically sensitive to olive pollen in southern Spain.

Pollen forecasting has become a major goal in aerobiology, and a number of papers specifically address aspects of olive pollen forecasting. Most authors agree that temperature is one of the main factors affecting the start of flowering in Mediterranean trees (Alba and Díaz de la Guardia, 1998; Frenguelli et al., 1989; Fornaciari et al., 1998; González-Minero and Candau Fernández-Mensaque, 1996; Galán et al., 2000). This is the case of *Olea europaea* L., which flowers from April to June in this area, depending on topography (Fornaciari et al., 2000) and variety. In spring-flowering trees, a low-temperature period of dormancy is required prior

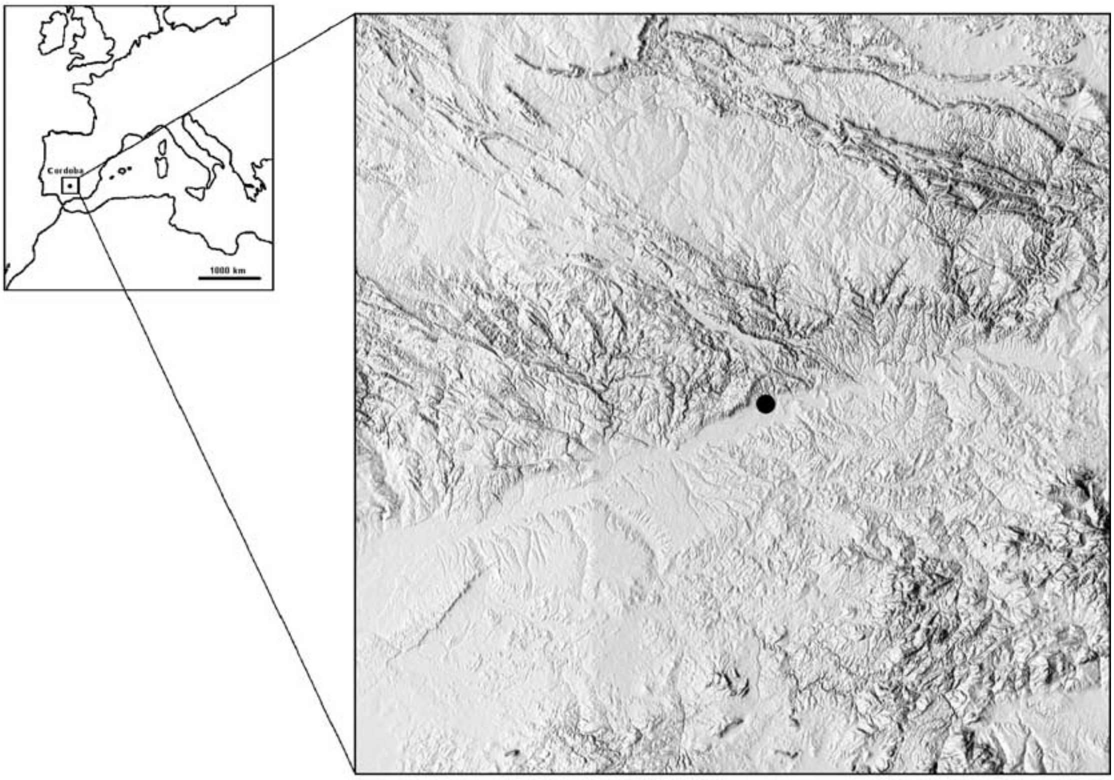

Figure 1. Location of sampling site, Cordoba, Spain. A 200 × 200 km square centred on the city of Cordoba.

to bud development. Development of reproductive structures occurs in response to temperatures above a threshold value and continues until flowering starts. A previous study in Cordoba has established base temperature and heat accumulation in order to forecast the start of the olive pollen-season using phenological (Alcalá and Barranco, 1992) and aerobiological data (Galán et al., 2000). The most important factor affecting olive pollen production in the Mediterranean area is rainfall over previous months (Fornaciari et al., 1997; González-Minero et al., 1998; Galán et al., 2001).

This study is being performed within the framework of a European Project named ASTHMA (Advanced System of Teledetection for Healthcare Management of Asthma), whose ultimate objective is to provide accurate near-real-time information on aeroallergens and air quality to sensitive users on an individual basis and at specific locations, in order to help them optimise medication use and improve their quality of life. The purpose of this system is to forecast pollen emission by a known source in terms of both space and time. The results obtained will form the basis of a whole system designed to help patients, which will include information on other pollen taxa such as *Parietaria*, *Poaceae* and *Cupressaceae*, as well as other airborne contaminants and medical data.

Kawashima and Takahashi (1995) have developed a pioneering system to simulate the emission and dispersion of airborne pollen cedar in Japan. In 1999, they improved the system by including a flowering-time map which takes into account the effects of altitude and local climate. The idea of a potential pollen production map was described by Puppi Branzi and Zanotti (1992) for *Castanea*. The authors have highlighted the difficulties involved in producing a detailed forecasting model which reflects real pollen production per unit area.

The main goal of this study was to develop a forecasting system comprising an emission model, a meteorological model and a dispersion model. This system is intended to produce near-real-time concentration maps for selected allergens. The present report on automated olive pollen forecasting is a preliminary step.

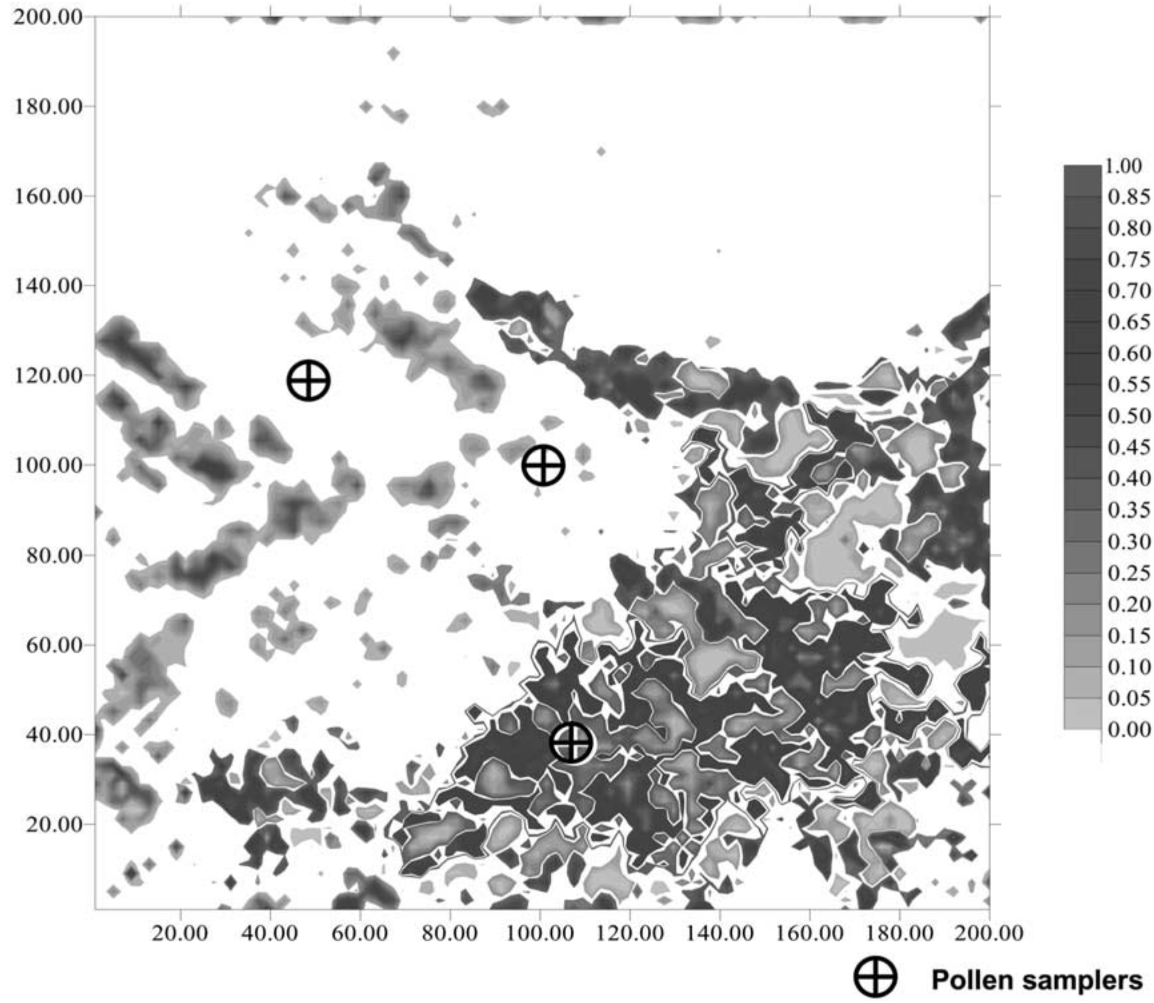

Figure 2. Pollen Production Map (PPM). A data matrix of 200x200 km centred on the city of Cordoba with a resolution of 1 × 1 km. Total surface area covered by olive trees was used in producing this data matrix, giving a maximum value of 1 when the full surface was covered.

Input data

System input data were split into three categories: static data, dynamic data and calibration parameters.

1. **Static data:** These data do not vary over a short time period. They can be divided into:
   - 1.1 Topographical data: Topographical data were extracted from the global GTOPO30 database.
   - 1.2 Land use data: High resolution satellite data provide information on the location of olive trees in the area of interest. A land-use and floral-spectrum digital database of Andalusia (source SINAMBA Sensor TM, resolution 1:50.000) was used as input. This database includes specific codes for land cover uses and floral spectra; the specific code for olive trees was selected. This information was translated into a 200 × 200 km data matrix centred on the city of Cordoba, with a resolution of 1 × 1 km. The total surface area covered by olive trees was taken into account when producing this data matrix, giving a maximum value of 1 when the full surface was covered. The matrix was used to generate the Pollen Production Map (PPM) with values between 0 and 1 (Figure 2). This PPM will be the basis of the emission module.
   - 1.3 Flowering statistics: statistics for flowering are required for computation and preparation of the flowering forecast. This information can be found in archived data from the aerobiological sampler and in field phenological observations.

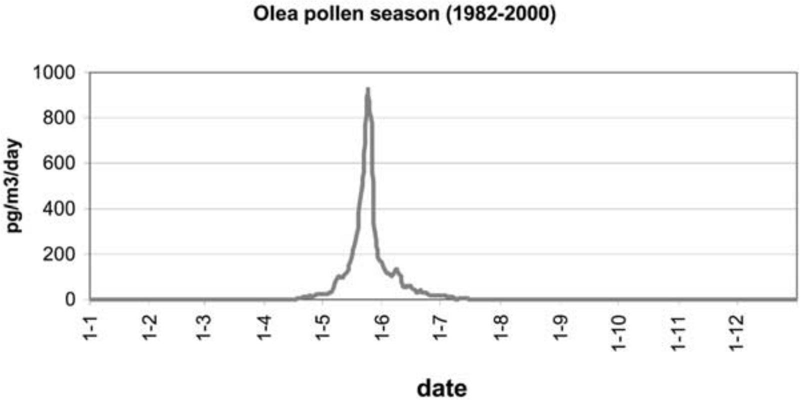

Figure 3. Peak date aligned average pollen concentration for the *Olea* pollen season from 1982 to 2000 for Cordoba, Spain.

Figure 3 shows the peak date aligned average pollen concentration for the *Olea* pollen season from 1982 to 2000 for Cordoba, Spain. Aerobiological data were collected using a Hirst-type spore-trap, located on the roof of the Faculty of Science, at 15 m above ground level. Standard sampling procedures proposed by the Spanish Aerobiology Network (REA) were used (Dominguez et al., 1992). The start of the *Olea* pollen season has been defined as the date on which 1 pollen grain/m^3 was recorded, as long as it was followed by at least 5 consecutive days also recording 1 or more pollen grains/m^3 (Galán et al., 2000). Field phenological observations were also carried out in order to enhance the aerobiological forecast. These observations were necessary not only for the preparation of a model but also for validation of the system.

2. **Calibration parameters:** these parameters are used to adjust the system during emission and dispersion once the static data have been established. Four parameters are considered:

2.1 Pollen season characteristics. The thermal summation method, using a threshold temperature of 10.4 °C, yields a result of 255.4 GDD° required for development prior to the start of the pollen season. Neural networks have been applied to forecast severity (total pollen count during the pollen season), using as inputs historical pollen data and meteorological parameters for previous months.

2.2 Dispersion factors: Dispersion factors include the amount of pollen produced by the plant, and the effect of meteorological conditions on dispersion. The location and distance of *Olea* crops were also considered. Density and intrinsic characteristics of *Olea* crops should always be taken into account in order to enhance the forecast. Other factors affecting dispersion, such as topography, cities, forest, etc. were also reflected. A Relative Production Factor (RPF) covering all dispersion factors and pollen production will be defined for each cell of the PPM. The relative value of the PPM (between 0 and 1) is multiplied by the RPF in order to obtain a relative value for pollen concentration.

2.3 Pollen production: total pollen production per tree was calculated in order to convert PPM relative emission to real emission. Total pollen grains per tree were counted using the method proposed by Hidalgo et al. (1999). Total number of flowers per square meter was estimated in 40 individuals by counting the number of flowers/raceme (11.1 ± 1.2), the number of racemes/branch (10.2 ± 2.2) and the number of branches/m^3 (74.5 ± 8.7). Total surface area was estimated by considering the tree to be spherical in shape and estimating the average radius (2.0 ± 0.3). The number of flowers per tree can be calculated by multiplying the number of flowers per m^2 by the total surface of the tree. The result was 107,394 flowers per tree. Total amount of flowers per tree may vary from year to year. This value will be calculated again in following years in order to ascertain how meteorological parameters might affect flower production. Similarly, the total number of pollen grains per flower was estimated in

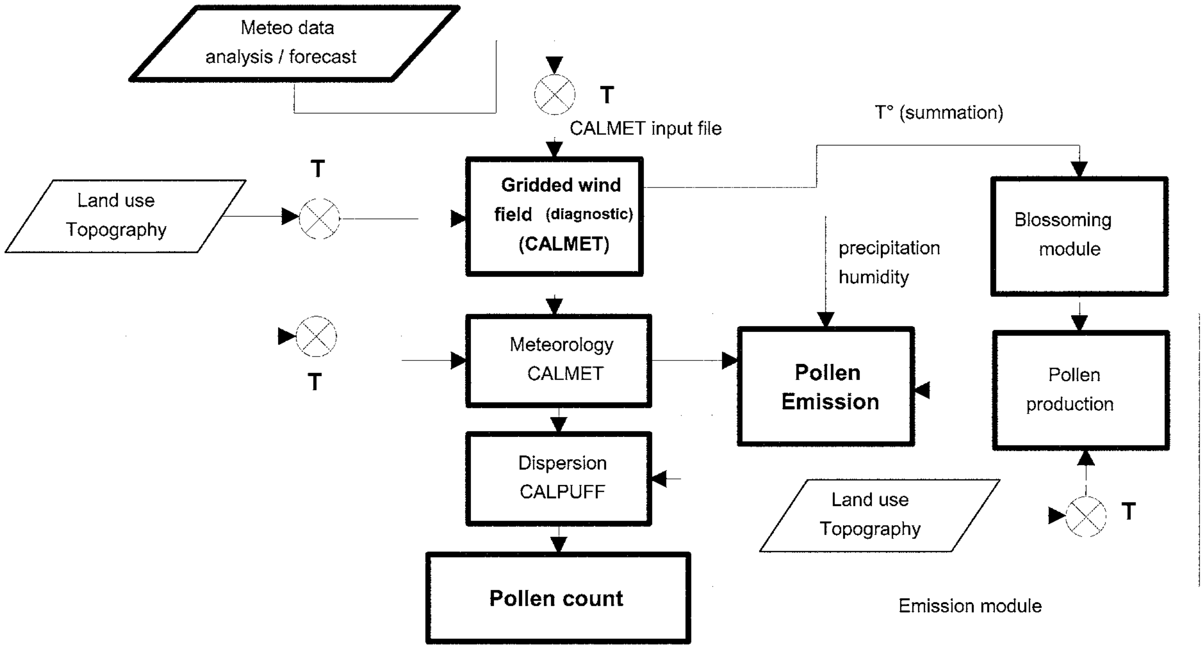

Figure 4. Architecture of the system. Inputs, computation modules and outputs.

30 flowers by crushing fresh anthers in 5 ml of distilled water. Total pollen was counted from this concentrate by depositing 10 μl on a slide. The result was 188,106 $\pm$ 8,554 pollen grains per flower. Total pollen grain production per tree was 20.2 million. Average tree density was also calculated by measuring the distance between trees. The result was 8.5 $\pm$ 1.4 m (based on 28 sampling sites) and the total number of trees per square km was 13,840.

2.4 Emission factors: Effective pollen emission from the plant is by far the most difficult parameter to quantify, and is still under study. Some basic methods are currently being tested for obtaining a convenient emission factor. The main idea is to minimise the error between measurement and forecasting by applying a uniform emission factor at a given time that is not wind-dependent. The correlation of the fluctuation of this coefficient with wind variations will be examined in order to obtain a suitable ratio. The main parameters affecting emission factors are the physiology of the plant and certain meteorological features including relative humidity, temperature, rainfall, etc.

3. **Dynamic data:** these data are required for the computational modules. Their main characteristic is the considerable variability recorded over a short period of time. Data can be classified as follows:

3.1 Meteorological data: Archive and real-time meteorological data from the pilot site are used. For the meteorological module, data from the COAMPS (Coupled Ocean Atmospheric Mesoscale Prediction System) model at Monterey University (California) have been used.

3.2 Pollen data: Archive and real-time data from different pollen traps at the pilot site are expressed as daily average pollen grains per cubic meter. Three pollen samplers are available in the study area (Figure 2). At present we are forecasting with the data available from the central pollen sampler (placed on the roof the Faculty of Science, University of Cordoba, Spain).

Computational modules

Figure 4 shows the architecture of the system. For computation, three different modules are involved: emission, meteorological and dispersion modules. Meteorological and dispersion modules are governed by the CALPUFF model. The CALPUFF Modeling System comprises three basic components: CALMET,

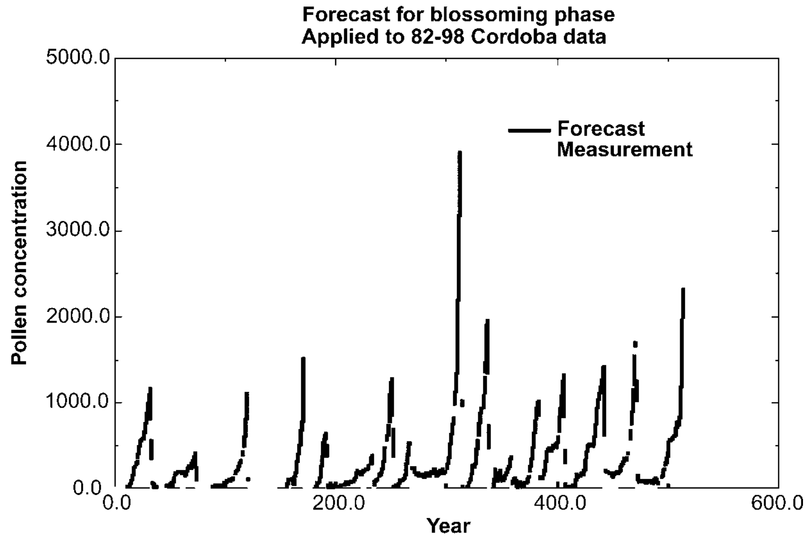

Figure 5. Forecasting and measuring pollen index (data from 1982 to 1998). Each peak represents one year (arbitrary units). The coefficient of correlation between the two pollen curves was 0.88.

CALPUFF and CALPOST. CALMET is a meteorological processor which was modified to enable use of vertical wind and temperature profiles as characterized by the MM4-FDDA (Mesoscale Model-4 with Four-Dimensional Data Assimilation) meteorological model. CALPUFF is a Lagrangian puff dispersion model and CALPOST is a postprocessor program that has been modified for ease of use and interpretation. CALPOST was modified to include a light extinction algorithm for use in regional visibility impact assessments.

1. *Emission module*

The emission module computes the pollen emission rate in every cell of the computation domain. This computation can be split into three categories:

- Spatial characterisation: PPM with values from 0 to 1
- Temporal characterisation: Start of pollen season (thermal summation), severity and daily pollen forecast (neural networks).
- Absolute quantification: multiplying by nominal pollen production per unit area.

2. *Meteorological module*

The meteorological module makes use of data coming from low-resolution meteorological analysis, such as wind field, temperature, pressure, atmospheric moisture and rainfall rates. The meteorological module is activated using a meteorological diagnostic model which interpolates (according to mechanical and thermal considerations) available data on a regular high-resolution data grid. The diagnostic model used is the CALMET model developed by Earth Tech, Inc.

3. *Dispersion module*

The dispersion module is the core of the forecast system. From the initial pollen concentration maps, the meteorological data and the emission maps, the dispersion module computes predicted concentration maps. The dispersion module is effected using the CALPUFF model (Earth Tech), which can be adjusted to reflect sedimentation laws, rain wash and thermal effects.

Forecast

The main goal is to forecast the time patterns of pollen production, emission and dispersion at the pilot site. The difficulty resides in deriving phenological data from aerobiological data. Rapid fluctuations in pollen concentration are due to dynamic and external factors (aerodynamic factors). Consequently, many other factors such as rainfall, temperature, etc. are not taken into account when the phenological status is obtained by smoothed pollen-count data.

The best way to forecast pollen concentration is to consider data from today (D) in order to fore-

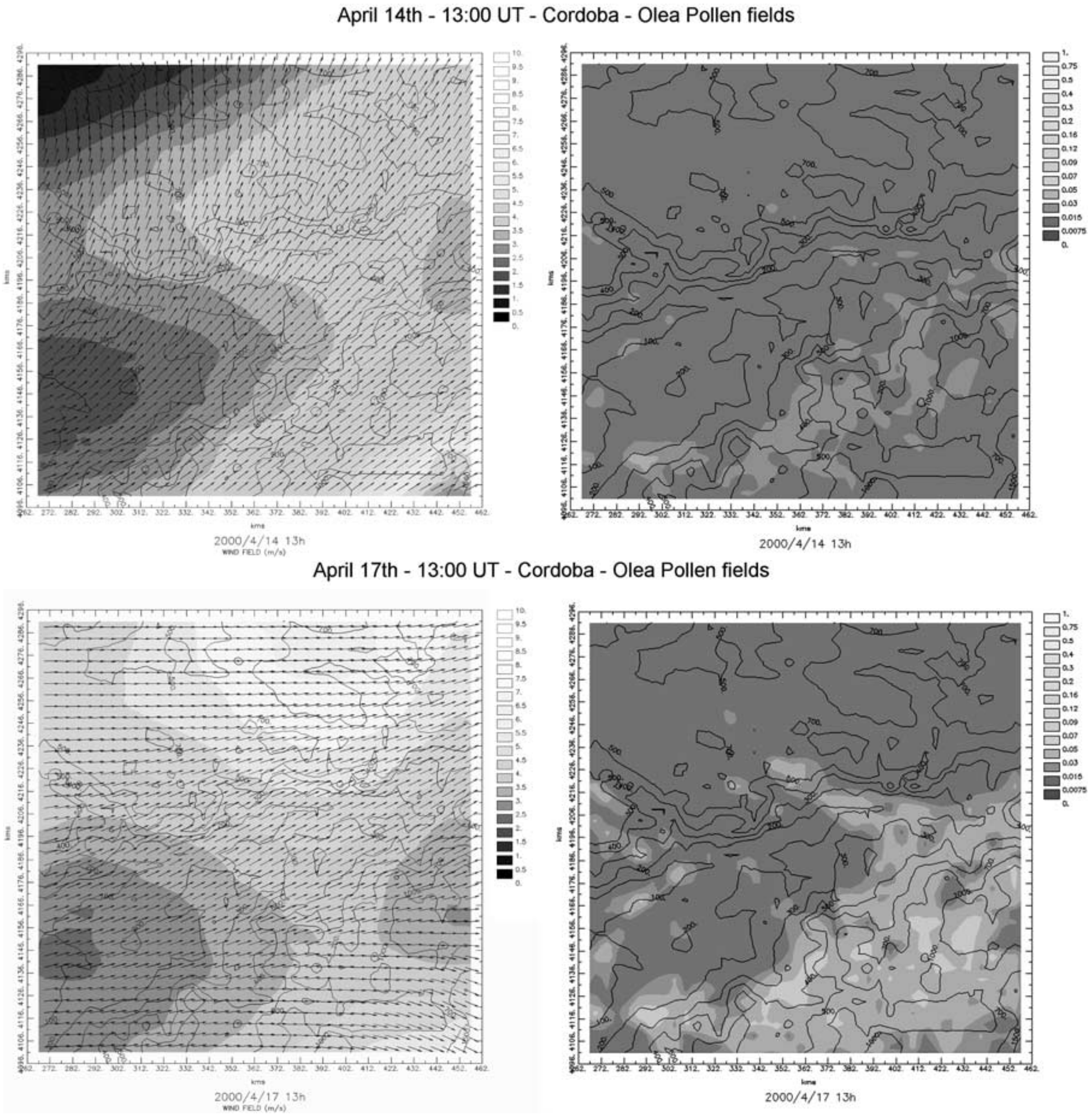

Figure 6. Simulation of wind field and pollen concentration. This simulation comprises one week from the start of the pollen season (April 14th, 17th and 20st).

cast the concentration for tomorrow (D+1). Forecasting for the following day (D+1) is performed using neural networks. Inputs are: maximum, minimum and average temperature, rainfall on D, sum of temperatures, sum of rainfall, sum of pollen from 1st January up to D, and an estimation of annual severity. Figure 5 shows observed versus expected values of daily average pollen for the years 1982 to 1998, using this method. An association between observed and expected values has been observed, despite some bad years. The coefficient of correlation between the two curves is 0.88. The poor matching of the curves in bad years may be due to other external factors such as wind (speed and modulus), which are not taken into account in the neural network analysis. In any case, the meteorological module includes these variables in order to improve the forecast during the simulation fields (see Figure 6).

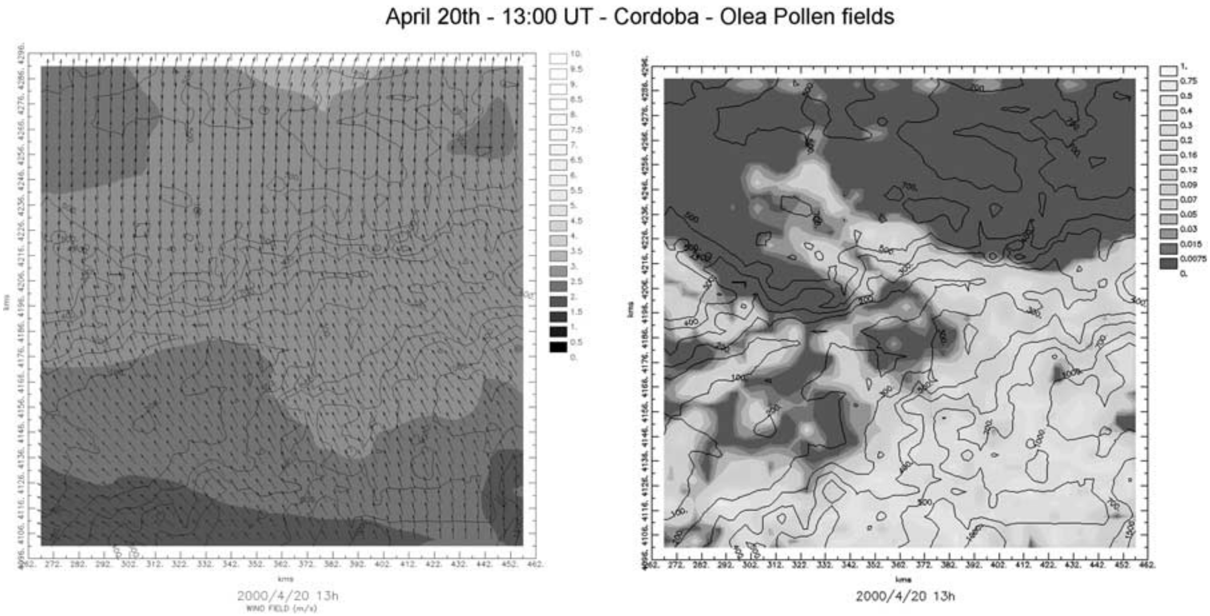

Figure 6. Continued.

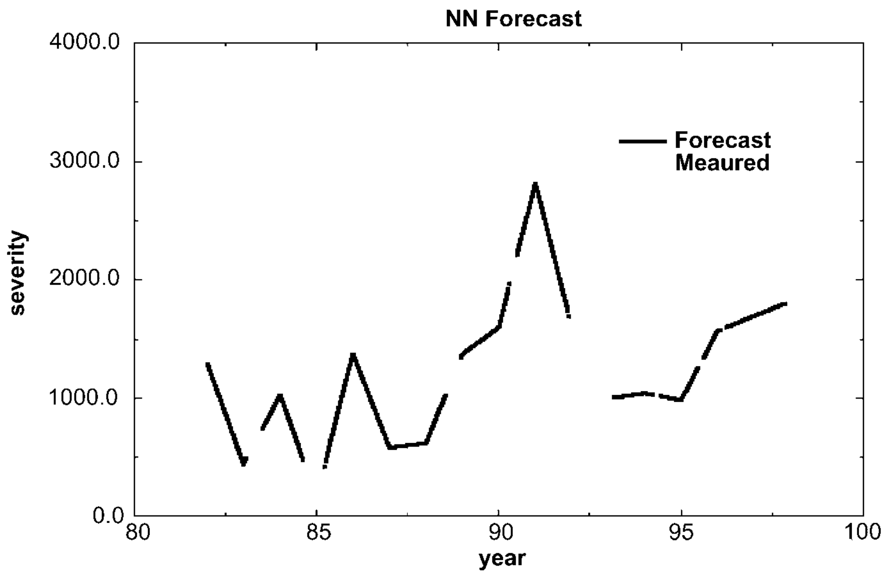

Figure 7. Neural network forecasting of severity during study years (1982 to 1998). The coefficient of correlation between the two curves was 0.92.

Severity (total pollen count during the pollen season) can be forecasted using neural networks. Figure 7 shows results for observed versus expected values for severity from the period 1982–98. In general, the same trend was observed in both curves. Total *Olea* pollen count in 2000 was 10,703 pollen grains, compared to a forecast of 11,100 pollen grains using neural networks.

Simulation

A simulation of wind field and pollen concentration is shown in Figure 6. This simulation comprises one week from the start of the pollen season (14th April). Pollen concentration fields are not validated and the concentration of pollen clouds should be multiplied by the relative production factor in order to obtain the real concentration (the concentration has been normalised here). The effect of topography has not yet been taken into account in this simulation of phenological behaviour, as indicated in the introduction. The emission factor has been set at 1.

Conclusions

A highly-innovative method has been developed to forecast *Olea* pollen in the atmosphere, and will be validated next spring. However, from the results obtained so far, the neural network analysis would appear to be a good tool for forecasting the start of the *Olea* pollen season, its severity and daily average pollen counts. Nevertheless, more variables, such as wind speed and modulus, are needed to enhance the method.

The temporal distribution of pollen release from a single tree as a function of external factors (meteorology) and internal factors (biology) should be considered and is still under investigation. A theoretical pollen-release curve should be established during the validation period in order to characterise this temporal distribution.

Pollen data from other samplers at the pilot site, i.e. data from the North and South, could be useful when adjusting the varying phenological behaviour of different tree populations. By this means, it will be possible to chart pollen emission patterns not only in space but also in time.

Acknowlegements

The authors are grateful to the European Community, project ENV4-CT98-0755 and to ZAMBON Group Company for financial support.

References

Alba, F. and Díaz de la Guardia, C.: 1998, The effect of air temperature on the starting dates of the *Ulmus*, *Platanus* and *Olea* pollen season in the SE Iberian Peninsula. *Aerobiologia* **14**, 191–194.

Alcalá, A.R. and Barranco, D.: 1992, Prediction of flowering time in olive for the Cordoba Olive Collection. *HortScience* **27**(11), 1205–1207.

Puppi Branzi, G. and Zanotti, L.Z.: 1992. Estimate and mapping of the activity of airborne pollen sources. *Aerobiologia* **8**, 69–74.

Domínguez, E., Galán, C., Villamandos, F. and Infante, F.: 1992, Handling and evaluation of the data from the aerobiological sampling. Monografías REA/EAN N° 1. Editado por el Departamento de Biología Vegetal y Ecología, Universidad de Córdoba. (D.L.:CO-476-1992).

Domínguez, E., Infante, F., Galán, C., Guerra, F. and Villamandos, F.: 1993, Variation in the concentrations of airborne *Olea* pollen and associated pollinosis in Cordoba (Spain); a study of the ten years period 1982–1991. *Journal of Investigational Allergology & Clinical Immunology* **3**(3), 121–129.

Florido, J.F., González Delgado, P., Saenz de San Pedro, B., Quiralte, J., Arias de Saavedra, J.M., Peralta, M. and Ruiz Valenzuela, L.: 1999, High levels of *Olea europaea* pollen and relation with clinical findings. *Int. Arch Allergy Immunol.* **119**, 133–137.

Fornaciari, M., Galán, C., Mediavilla, A., Domínguez, E. and Romano, B.: 2000, Aeropalinological and phenological study in two different olive Mediterranean areas. *Plant Byosistem* **134**(2), 199–204.

Fornaciari, M., Pieroni, L., Ciuchi, P. and Romano, B.: 1998, A regression model for the start of the pollen season in *Olea europaea*. *Grana* **37**, 110–113.

Frenguelli, G., Bricchi, E., Romano, B., Mincigrucci, G. and Spieksma, F.T.H.M.: 1989, Predictive study on the beginning of the pollen season for Gramineae and *Olea europaea* L. *Aerobiologia* **5**, 64–70.

Galán, C., Cariñanos, P., Alcázar, P., García-Mozo, H. and Domínguez, E.: 2000, Forecasting model to predict the *Olea europaea* L. airborne pollen production in South West Andalusia, Spain. (submitted to publication).

Galán, C., Garcia-Mozo, H., Cariñanos, P., Alcázar, P. and Domínguez, E.: 2001, The role of temperature in the onset of the *Olea europaea* L. pollen season in South-western Spain. *Int. Journal Biometeorology* (in press).

González-Minero, F.J. and Candau Fernández-Mensaque, P.: 1996, Prediction of the beginning of the olive full pollen season in south-west Spain. *Aerobiologia* **12**, 91–96.

González-Minero, F.J., Candau, P., Morales, J. and Tomas, C.: 1998, Forecasting olive production based on ten consecutive years of monitoring airborne pollen in Andalusia (Southern Spain). *Agriculture, Ecosystems and Environment* **69**, 201–215.

Hidalgo, P.J., Galán, C. and Domínguez, E.: 1999. Pollen production of the genus *Cupressus*. *Grana* **38**, 296–300.

Kawashima, S. and Takahashi, Y.: 1995. Modelling and simulation of mesoscale dispersion processes for airborne cedar pollen. *Grana* **34**, 142–150.

Kawashima, S. and Takahashi, Y.: 1999, An improved simulation of mesoscale dispersion of airborne cedar pollen using a flowering-time map. *Grana* **38**, 316–324.