

This article was downloaded by:[Advanced Institute of Sanita]
On: 9 February 2008
Access Details: [subscription number 769790694]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



International Journal of Environmental Health Research

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title-content=t713425582>

Modelling atmospheric concentrations of grass pollen using meteorological variables in Melbourne, Australia

Bircan Erbas ^a, Jiun-Horng Chang ^a, Ed Newbigin ^b, Shyamali Dhamarge ^a

^a Centre for Molecular Environmental Genetic Analytic Epidemiology, School of Population Health, University of Melbourne, Victoria, Australia,

^b Plant Cell Biology Research Centre, School of Botany, The University of Melbourne, Victoria, Australia

Online Publication Date: 01 October 2007

To cite this Article: Erbas, Bircan, Chang, Jiun-Horng, Newbigin, Ed and Dhamarge, Shyamali (2007) 'Modelling atmospheric concentrations of grass pollen

using meteorological variables in Melbourne, Australia', International Journal of Environmental Health Research, 17:5, 361 - 368

To link to this article: DOI: 10.1080/09603120701628693

URL: <http://dx.doi.org/10.1080/09603120701628693>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Modelling atmospheric concentrations of grass pollen using meteorological variables in Melbourne, Australia

BIRCAN ERBAS¹, JIUN-HORNG CHANG¹, ED NEWBIGIN², &
SHYAMALI DHAMARGE¹

¹Centre for Molecular Environmental Genetic Analytic Epidemiology, School of Population Health, University of Melbourne, Victoria, Australia, and ²Plant Cell Biology Research Centre, School of Botany, The University of Melbourne, Victoria, Australia

Abstract

This study aims to develop models that may be used to describe the relationship between meteorological variables and ambient concentrations of pollen. We used daily ambient concentrations of grass pollen during the pollen season (October to December) of 2004 in Melbourne, Australia. During this period, daily levels of meteorological data including average relative humidity, mean temperature, rain fall, wind speed, and wind direction were entered as predictors in the models. A generalized additive model (GAM) was used to assess the relationship between daily levels of meteorological variables and ambient concentrations of grass pollen. The relationship between average temperature, rain fall, wind speed, relative humidity and pollen were nonlinear and smooth terms were highly significant ($p < 0.001$). Nonlinear statistical methods such as the GAM approach have the potential to accurately predict ambient concentrations of pollen during the pollen season.

Keywords: *Ambient pollen, meteorological variables, statistical models, nonlinearity*

Introduction

In many countries of diverse populations and climate; individuals, particularly children are susceptible to pollen induced allergens in both chronic and acute forms. In fact, most asthmatics are atopic to common aeroallergens including grass pollens. Increased hospitalizations due to asthma have been observed in high pollen seasons (Tobias et al. 2003). Short-term predictions of atmospheric pollen concentrations during the pollen season are essential for informing health professionals and the general community of impending days of high concentrations of ambient pollen concentrations.

For a number of decades, aerobiological research have focused on formulating mathematical models and applying varying statistical methods to predict both short- and long-term daily airborne pollen concentrations (Jato et al. 2000; Laaidi et al. 2003; Cotos-Yanez et al. 2004; Castellano-Mendez et al. 2005; Smith 2005). Developing models to

Correspondence: Dr Bircan Erbas, Centre for Molecular Environmental Genetic Analytic Epidemiology, School of Population Health, University of Melbourne, Level 2, 723 Swanston Street, Carlton 3053, Australia. Tel: +613 8344 0635. Fax: +613 9349 5815. E-mail: b.eras@unimelb.edu.au

predict daily pollen concentrations during the pollen season is rather complex and poses a number of key methodological constraints. Both pollen and meteorological data are time series where daily successive observations are correlated with each other. Moreover, daily trend and the relationship between pollen and each of the meteorological variables may be linear, nonlinear or both linear and nonlinear captured via semi-parametric estimates.

The predictive power of any model is highly dependent on adequately modelling the variation in daily ambient concentrations of pollen. Many standard time series techniques (Box-Jenkins methods, Transfer Function Models) are inappropriate for investigating these relations because either the methodological assumptions are inadequate or the methods are difficult to implement. The statistical approach adopted in studies attempting to produce short- or long-term pollen forecasts have been correlations (Jato et al. 2000; Green et al. 2003) linear and logistic regression models (Laaidi et al. 2003; Cotos-Yanez et al. 2004; Rodriguez-Rajo et al. 2004; Smith 2005). All studies incorporated linear effects of meteorological variables. Others have used a neural networks approach to capture nonlinear relationships with temperature (Ranzi et al. 2003; Castellano-Mendez et al. 2005). However, the 'black box input-output' approach in neural networks is limited by a lack of interpretation and unknown relevance for public health policy.

Recently, Cotos-Yanez and colleagues (2004) used logistic additive models to predict short-term concentrations of *Betula* airborne pollen. Although, additive models have been widely adopted in other areas of science they are yet to be applied to data in aerobiological research. The major advantage of this approach is that they allow the pollen concentrations to vary smoothly as a function of meteorological variables. A graph of the fitted curves and 95% pointwise confidence intervals allow a simple visual interpretation and the models can be applied with relative ease in most standard statistical packages. In this study, we use generalized additive models to model variation in daily concentrations of ambient grass pollen and smooth terms for meteorological variables to assess the adequacy of these techniques for predictive purpose in future studies.

Methods

Setting

This was an ecological study conducted using meteorological data and daily pollen counts from late September through to the end of December, a period that covers the entire grass pollen season in Melbourne, Australia (Ong et al. 1995).

Melbourne (37:48:09.60S 144:57:29.50E) is the second largest city in Australia and is situated on Port Phillip Bay (Figure 1). The climate is generally temperate with warm summers and cool winters. The mean daily maximum temperature ranges from 26°C in January, the hottest month, to 13°C in July, the coldest month. Hot days in Melbourne are due to hot winds from the north and north-west. During the grass flowering season, these winds bring in pollen from pasture grasses growing in western Victoria, the most common pasture grass being perennial rye (*Lolium perenne*). Cooler winds from the south-east are also responsible for bringing pollen into the city, but winds from other directions do not bear a significant pollen load (Smart & Knox 1979).

Data

Pollen data. A Burkard volumetric trap fitted with a 24-h sampling head (Burkard Scientific Ltd, Middlesex, UK) was used for the pollen counting. Airborne grass pollen was counted on

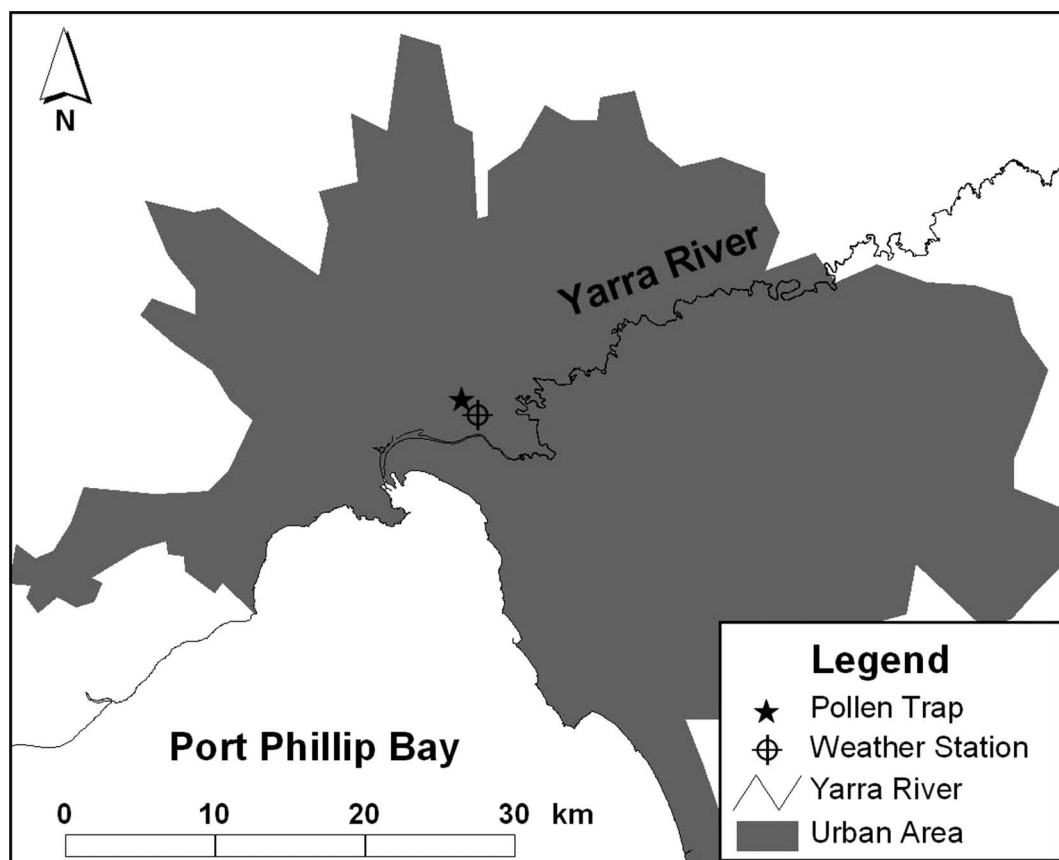


Figure 1. Location of pollen trap and weather monitoring stations in Melbourne, Victoria, Australia.

a daily basis from late September through to the end of December. The location of the trap, on the city campus of the University of Melbourne, conforms to the Australian Standard AS 2922-1987 for the siting of ambient air sampling units for environmental monitoring (Standards Association of Australia, Sydney, NSW). Pollen grains were trapped on a glass microscope slide ($76.2 \times 25.4 \times 1.2$ mm) coated with an adhesive (Dow Corning Sylgard 527 silicone dielectric gel) by intake of air at 10 l min^{-1} . Each day during the sampling period, the slide, attached to a rotating drum on the sampling head, was removed at around 3 pm and replaced with a fresh slide. The daily concentration of grass pollen (in grains per m^3) was determined by staining the slide with Calberla's stain (Odgden et al. 1974) and counting grass pollen grains along a random traverse of the slide's length and averaging this value across the entire slide.

Meteorological data. Meteorological data (including daily rainfall, % relative humidity, hours of sunshine, maximum, minimum, and mean temperature) were provided by Australian Bureau of Meteorology from data collected at a weather station 1 km from the pollen collection site. Wind speed (m/s) and wind direction classified as north (N), northeast (NE), east (E), southeast (SE), south (S), southwest (SW), west (W) and northwest (NW) were included as potential predictors in the models. Daily rain fall (mm) was also entered as a continuous predictor.

Statistical methods

The response variable pollen is a count and thus was assumed to be Poisson distributed. Graphical methods such as scatter plots with smooth curves fitted by Lowess smooths were used to explore trends and potential nonlinearity between pollen counts and each of the explanatory variables. In the multivariate models explanatory variables were entered as days in pollen season (1 October 2004 – 31 December 2004), rainfall as a continuous variable, average temperature, wind speed, wind direction and average humidity.

Correlations between each of the explanatory variables were checked for multicollinearity before fitting the model with the largest correlation coefficient between average temperature and average humidity; rain and average humidity. However, all estimated correlation coefficients were less than $|0.5|$.

A Poisson regression was also specified in Generalized Additive Model (GAM) (Hastie & Tibshirani 1990) to predict pollen counts using nonparametric penalized smoothing splines with smoothing parameters automatically determined by Generalized Cross-Validation procedure (GCV), and using chosen fixed degrees of freedom for the smoothing terms where adjustments of smoothness was desired. Included in this model was a parametric term for wind direction, and five continuous variables (days, average temperature, rainfall, wind speed, and average humidity) all modelled with smooth parameters. Diagnostic plots were checked for model fit and distribution of residuals. No serial correlations were shown in autocorrelation plot (ACF plot) in the models. All statistical tests were two-sided with significance level chosen at 5%. All statistical analyses were performed in R implementation of the S language (Version 2.0.1).

Results

Summary measures for grass pollen and meteorological variables during October through to December in 2004 are presented in Table I. During the grass pollen season of 2004, 51 days were low grass pollen days (≤ 20 grains/m³), 18 were moderate grass pollen days (21–50 grains/m³), 11 were high grass pollen days (50–100 grains/m³), and 11 days were extreme grass pollen days (> 100 grains/m³). Both exploratory and univariate analyses of all continuous variables exhibited some nonlinear relationship with pollen counts.

Results from fitting a GAM with log link are displayed in Table II. In the multivariable additive model, where all significant meteorological variables are entered into a model, all nonparametric smooth components were highly significant suggesting a nonlinear relationship between the meteorological variables and pollen. The fitted values from the GAM of pollen and meteorological variables in Figure 2 display a reasonably adequate fit to the pollen data.

Table I. Summary measures of grass pollen and meteorological variables during the 2004 pollen season.

	Min	1st Quartile	Median	Mean (SE) ^a	SD ^a	3rd Quartile	Max
Grass pollen	0	8	18	43.64	–	47.50	317
Rainfall – overall	0	0	0.00	2.69	–	1.7	26
– rainy day ¹	2	1.5	3.8	7.73	–	11.3	26
Temperature (mean)	11.2	15.1	17.4	17.9 (0.38)	3.67	20.5	28.6
Wind speed	2.1	4.6	6.2	6.38 (0.26)	2.44	8.2	14.9
Humidity (mean)	27.25	54.59	60.94	61.80 (1.07)	10.24	68.69	87.50

¹During the pollen season in the year 2004: 32 days were wet days (rainfall > 0.05 ml) and 60 dry days (rainfall < 0.05 ml); ^acalculation based on data with outliers removed (if any).

Discussion

An accurate daily prediction of ambient concentrations of pollen is an essential component of public health policy used to inform health professionals and persons who are already allergic or susceptible to environmental allergies, particularly pollen grains in the atmosphere. In many cities with increased levels of atmospheric pollen, daily predictions are available and for

Table II. Results of Generalized Additive Model (GAM) of daily levels of pollen and meteorological variables during the pollen season of 2004.

Smooth terms	Degrees of freedom	Chi sq	<i>p</i> value
Days	9.00	336.8	<0.001
Average daily temperature	9.00	323.1	<0.001
Daily rainfall	8.94	176.9	<0.001
Wind speed	8.72	244.4	<0.001
Average daily humidity	8.71	250.1	<0.001
<i>Parametric terms:</i>			
	<i>p</i> value		
Wind direction	<0.001		
Rain indicator	<0.001		

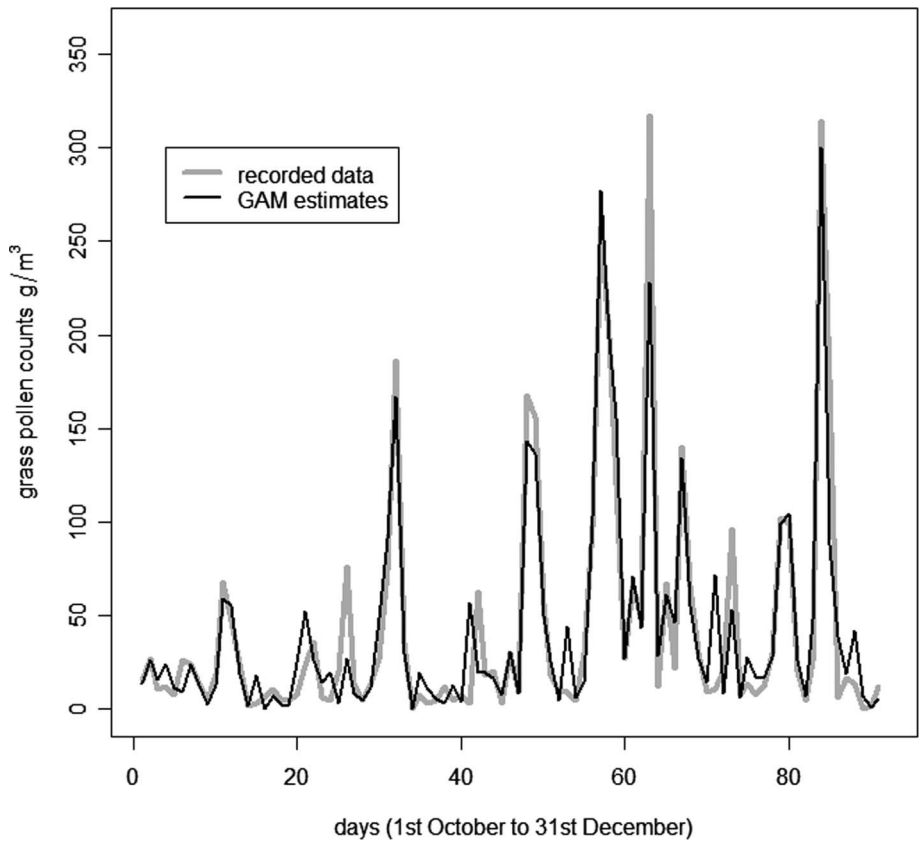


Figure 2. Fitted values from a Generalized Additive Model (GAM) with nonparametric smoothing terms overlaid on daily pollen levels during the pollen season of 2004.

some the current practice is at best a *good guess* approach with very little modelling beforehand. Although the end point, i.e. predicting ambient concentrations of pollen on a daily basis is important, a further understanding of the underlying trends in ambient concentrations of pollen and to adequately model this behaviour using appropriate analytical tools is fundamental to making accurate predictions. If we can adequately describe variations in ambient levels of pollen then these models can then be implemented to provide accurate estimates of pollen predictions.

Our data reflect the usual weather pattern and pollen counts in Melbourne (Ong et al. 1995). In this analysis of ambient levels of pollen during the pollen season of 2004 in Melbourne, Australia we present an alternative modelling framework for estimating ambient concentrations of pollen. Under this modelling framework, we estimate significant nonlinear relationships between daily levels of meteorological variables and pollen measured as a count.

This study uses the same additive methodological approach adopted by Cotos-Yanez and colleagues (2004) by estimating nonlinear effects for the meteorological variables and assuming a numeric outcome measure for ambient concentrations of pollen rather than the binary threshold variable used in their study. Cotos-Yanez and colleagues (2004) build on earlier linear logistic models by including the pollen variable as explanatory terms in the models. They suggest that estimating this additional parameter improves the low predictive power of these models considerably compared to earlier models with only meteorological variables as explanatory variables.

Our approach is relatively similar in that rather than pollen, we include nonlinear effects for meteorological variables that explained more of the residual variation in pollen concentrations, thereby greatly improving the predictive power of these models in future studies. In contrast to others, we do not incorporate pollen data from a number of years and only focus on data from the one year. However, when we fitted similar GAM models to pollen data for 1996–2003 separately, we found that the fitted models suggest similar nonlinear associations (data not shown) and the utility of these models are replicable over different data sets.

The association between meteorological variables and ambient concentrations of pollen demonstrated in our study are consistent with other studies. Correlation studies of the atmospheric concentrations of pollen during the pollen season have shown the effects of temperature and rainfall on *Alnus* pollen from Spain (Rodriguez-Rajo et al. 2004), *Pinus* pollen in Brisbane, Australia (Green et al. 2003) and humidity and rainfall on *Pinus* pollen in the city of Vigo (Jato et al. 2000). Likewise, Vazquez and colleagues (2003) consistently showed effects of daily humidity on *Olea* pollen concentrations during their south-western Spain via linear regression models separately for most years between 1982 and 2002.

Damialis and colleagues (2005) report strong associations between wind direction, wind speed and airborne pollen using factorial regression methods. In addition, the independent and interaction effect of wind direction and speed on ambient pollen was confirmed in our study. A recent large 15-year study of 16 types of pollen associated with respiratory allergic symptoms in Thessaloniki, Greece (Gioulekas et al. 2004) suggest strong positive associations between wind speed and 8 different types of pollen. The scatter plot in figure D (Gioulekas et al. 2004) noticeably shows a nonlinear pollen–wind speed relationship and thus the low coefficient of variation in part may be explained by the incorrect assumption of linearity.

There are a number of strengths of the analytical approach adopted in this study. The GAM is effective exploratory and modelling techniques useful for modelling nonlinear effects. In contrast to the common simple correlation or linear regression approach used in many studies of predicting pollen using meteorological data, GAM techniques make no parametric

assumptions about meteorological variables. This reduces the potential of making incorrect inferences about the ambient pollen concentrations – meteorological variables relationships, thereby resulting in a model with greater precision and higher predictive power of pollen concentrations. Secondly, the GAM techniques accommodate for correlated errors a consequence of data collected consecutively over time, such as the case for ambient pollen concentrations during the pollen season.

It is important to recognize that any model for predictive purposes developed using a given data set may not be transportable to a new data set. In fact, even a good fit to the current data does not guarantee that the model will have predictive power (Royston 2000) particularly given the uncertainty in estimating meteorological data from year to year. Nevertheless, our findings provide a better understanding of the mechanism underlying trends in ambient pollen concentrations and its relationship with meteorological variables during the pollen season.

At present, during the pollen season the Australian Asthma foundation offers forecasts of a qualitative measure of pollen on a scale from low to extremely high (www.asthma.org.au). Given the likelihood of unexplained errors in any prediction model, the models used in this study are an improvement over current methods of prediction. Although, the models adequately describe the behaviour of pollen on a given day using meteorological variables on that day, refinements of the models are necessary to predict subsequent pollen counts. For example, reliable estimates of meteorological conditions are necessary in place of actual observations for these models to be used for predictive purposes.

A model with observed daily continuous values of pollen as the response variable predicts a daily continuous measure with associated error. A daily predicted value above a specified threshold that may exacerbate symptoms in already allergic individuals is extremely useful in comparison to a qualitative value. This will ideally require quantifying individual exposures and correlating that with individual symptom status. Less rigorous but a more feasible option would be to correlate ecological data using the appropriate statistical techniques. Such a study will enable identification of such cut-off points that can be directly transferable to public health practice.

Acknowledgements

This research was supported by a Victorian Asthma Foundation research project grant.

References

- Damialis A, Gioulekas D, Lazopoulou C, Balafoutis C, Vokou D. 2005. Transport of airborne pollen into the city of Thessaloniki: the effects of wind direction, speed and persistence. *Int J Biometeorol* 49:139–145.
- Castellano-Mendez M, Aira MJ, Iglesias I, Jato V, Gonzalez-Manteiga W. 2005. Artificial neural networks as a useful tool to predict the risk level of *Betula* pollen in the air. *Int J Biometeorol* 49:310–316.
- Cotos-Yanez TR, Rodriguez-Rajo FJ, Jato MV. 2004. Short-term prediction of *Betula* airborne pollen concentration in Vigo (NW Spain) using logistic additive models and partially linear models. *Int J Biometeorol* 48:179–185.
- Gioulekas D, Balafoutis C, Damialis A, Papakosta D, Gioulekas G, Patakas D. 2004. Fifteen years' record of airborne allergenic pollen and meteorological parameters in Thessaloniki, Greece. *Int J Biometeorol* 48:128–136.
- Green BJ, Yli-Panula E, Dettmann M, Rutherford S, Simpson R. 2003. Airborne *Pinus* pollen in the atmosphere of Brisbane, Australia, and relationships with meteorological parameters. *Aeobiologica* 19:47–55.
- Hastie TJ, Tibshirani R. 1990. Generalized additive models. London; New York: Chapman and Hall.
- Jato MV, Rodriguez FJ, Seijo MC. 2000. *Pinus* pollen in the atmosphere of Vigo and its relationship to meteorological factors. *Int J Biometeorol* 43:147–153.
- Laaidi M, Thibaudon M, Besancenot JP. 2003. Two statistical approaches to forecasting the start and duration of the pollen season of *Ambrosia* in the area of Lyon (France). *Int J Biometeorol* 48: 65–73.

- Odgen EC, Raynor GS, Hayes JV, Lewis DM, Haines JH. 1974. Manual for sampling airborne pollen. New York: Hafner Press.
- Ong EK, Singh MB, Knox RB. 1995. Seasonal distribution of pollen in the atmosphere of Melbourne: An airborne pollen calendar. *Aerobiologia* 11:51–55.
- Ranzi A, Lauriola P, Marletto V, Zinoni F. 2003. Forecasting airborne pollen concentrations: Development of local models. *Aeobiologica* 1:39–45.
- Rodriguez-Rajo FJ, Dopazo A, Jato V. 2004. Environmental factors affecting the start of pollen season and concentrations of airborne *Alnus* pollen in two localities of Galicia (NW Spain). *Ann Agric Environ Med* 11:35–44.
- Royston P. 2000. A strategy for modelling the effect of a continuous covariate in medicine and epidemiology. *Stat Med* 19:1831–1847.
- Smart IJ, Knox RB. 1979. Aerobiology of grass pollen in the city atmosphere of Melbourne: quantitative analysis of seasonal and diurnal changes. *Aeobiologica* 11:51–55.
- Smith M, Emberlin J. 2005. Constructing a 7-day ahead forecast model for grass pollen at north London, United Kingdom. *Clin Exp Allergy* 35:1400–1406.
- Tobias A, Galan I, Banegas JR, Aranguiz E. 2003. Short term effects of airborne pollen concentrations on asthma epidemic. *Thorax* 58:708–710.
- Vazquez LM, Galan C, Dominguez-Vilches E. 2003. Influence of meteorological parameters on *Olea* pollen concentrations in Cordoba (south-western Spain). *Int J Biometeorol* 48:83–90.