

Modelling aerobiological time series. Application to Urticaceae

Jordina Belmonte^{1,*} & Miguel Canela²

¹*Unitat de Botànica, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, Spain;* ²*Departament de Matemàtica Aplicada i Anàlisi, Universitat de Barcelona, Spain*

(*author for correspondence)

Received 2 March 2001; accepted 16 September 2002

Key words: aerobiology, ARIMA models, autocorrelation, pollen dynamics, Friedman supersmoother, Gaussian mixture model, *Parietaria*, Spearman correlation, times series, *Urtica*

Abstract

This paper deals with the characterisation of aerobiological time series. The series discussed are annual series of mean daily concentration of Urticaceae pollen, obtained by means of Hirst traps on six sites of the Aerobiological Network of Catalonia (XAC). First, some alternatives for fitting a smooth trend to one of these series are discussed, and two satisfactory methods are presented, a Gaussian mixture model (parametric) and the Friedman supersmoother (non-parametric). Dividing daily concentrations by trend values, a series of residuals is obtained. The residual series can be characterised with simple conventional time series models, and some common patterns emerge from the parameter estimates. The long-term patterns of the pollen dynamics can be thus characterised by the trends, and the short-term behaviour by the residuals. The Spearman correlation is used in this paper as a measure of the similarity of two series.

Introduction

The aim of this paper is to present and discuss different alternatives for the characterisation of aerobiological time series and the identification of their main patterns. This paper deals with 1-year series of mean daily pollen concentrations, but some of its methods also work for longer series. We are interested in a characterisation presenting as much information as possible, but easy enough to be understood by non-specialist end users (in particular, by allergologists, pneumologists and allergic patients). Such a characterisation will be helpful when investigating the relationship between pollen and environmental factors.

This paper is focused on a single pollen type, Urticaceae. We have applied different methods to the pollen series of six sampling sites in the region of Catalonia, Spain, which are integrated in the Aerobiological Network of Catalonia (Xarxa Aerobiològica de Catalunya, XAC). Table 1 is a summary

of geographical and climatic characteristics of the sampling sites. In Catalonia, the Urticaceae family is composed of three genera (Bolòs and Vigo, 1989), *Urtica* (5 species), *Parietaria* (3 species), and *Soleirolia* (1 species). The pollen grains of these species cannot be distinguished under light microscopes (only *U. membranacea* can be clearly recognised), and thus, Urticaceae pollen data always refers to several species.

Some attempts have been done to apply time series analysis to pollen series. Comtois and Sherknies (1991) classify pollen types according to the shape of a polynomial trend fitted to daily pollen concentrations. Other work is devoted to relate the trend (Comtois and Sherknies, 1992) or the deviations with respect to the trend (Bringfelt et al., 1982; Moseholm et al., 1987) to meteorological factors. Nevertheless, although pollen data are naturally structured as time series, and the techniques of time series analysis are widespread, there has been no much progress in this line.

Aerobiological data

Urticaceae has been chosen because it is the most allergenic pollen in the Mediterranean area (see Belmonte et al., 1999), and presents a particular dynamics on the air along the year. The series discussed in this paper are series of daily mean pollen concentrations (pollen/m³), obtained by means of Hirst seven-day spore traps (Hirst, 1952). We followed the counting method proposed by Domínguez et al. (1991) for the Spanish Aerobiology Network (Red Española de Aerobiología, REA). Urticaceae pollen is present in the atmospheric spectrum almost every day, so that, for some of the methods presented in this paper, there is no need of splitting the series of daily concentrations into annual series, and one can take a single series, several years long, for each site.

All the XAC sampling sites are placed in urban environments, except for Bellaterra and Manresa, where the environment can be considered as urbanised or close to rural. The six localities have similar climate. Girona, Manresa and Bellaterra are more humid and fresh, and Barcelona and Tarragona hotter. Lleida has a more extreme climate, always dry, with cold winters and hot summers (Table 1). Urticaceae pollen is more abundant in temperate urban environments near the litoral, than in the inland and colder regions (Belmonte et al., 1999).

Statistical methods

If $C(t)$ is the value of a variable at time t , the set of the $C(t)$, for $t = 1, 2, \dots$, is a time series. In this paper, t denotes the day number, running from 1 to 365 (or 366), i.e. from January 1st to December 31st, for a 1-year series, and $C(t)$ is the mean daily concentration of Urticaceae pollen (pollen/m³). Usually, the first step in the analysis of a series is the decomposition of $C(t)$ as the sum of a trend value $T(t)$ and a residual, $E(t)$. This approach is useful, because the trend helps to identify the seasonal patterns of pollen data, and time series analysis techniques can be applied to the residual series. Moreover, pollen trends can be used to fill in the data gaps in aerobiological databases, due to malfunctioning of the trap or other causes.

Spearman correlation has been frequently used in pollen data analysis (e.g. in Belmonte et al., 1988), as it is not affected by extreme values, which are frequent in this type of data. In our study, we have used Spearman correlation values as a measure of the

similarity between series from different sites, for the same year. Our idea is that the similarity detected by a high Spearman correlation will lead to a similar characterisation, i.e. similar values of the parameters of the models fitted to the data. We also expect that the similarity in the pollen series can be related to coincidence in environmental factors, although this has been left for future work. Spearman correlation values are discussed in this paper, for the three series mentioned above, $C(t)$, $T(t)$, and $E(t)$. Multivariate techniques, like taxonomical dendrograms (Comtois and Sherknies, 1991), have not been used here.

Most techniques of time series analysis are valid only for stationary series. Stationarity includes the assumption that the series is a sequence of random variables with constant mean and variance. Such assumption is obviously unrealistic for pollen series, not only for $C(t)$, but also for $E(t) = C(t) - T(t)$, because the deviations from any smooth trend are higher (in absolute value) for higher trend values. For the series studied in this paper, we have found that, except for those days with $C(t)$ close to zero, the magnitude of the deviations is more or less proportional to the trend. In a different approach, Stark et al. (1997) assume (for *Ambrosia*) that the variance of the residuals is proportional to the trend, so that Poisson regression can be used. For the series of this study, a linear relationship between trend values and the residual standard deviation is more feasible than the one assumed by Stark et al. (1992).

Most time series analysis textbooks recommend a transformation of non-stationary series, such as a difference filter, a logarithm, or a square root, before fitting the trend. For pollen data, this has two disadvantages: a) in some cases the deviations are not stationary, even after transforming the data, and b) most transformations change the scale of the data, so that the interpretation of the models obtained becomes more difficult. Nevertheless, logarithmic transformations are frequently applied to pollen data (e.g. in Moseholm et al., 1987). A different approach, which has been adopted in this paper, is to use a multiplicative model, i.e. $C(t) = T(t)[1 + E(t)]$. An adequate choice of $T(t)$ leads to a series of residuals $E(t)$ with mean close to zero and roughly constant variance. For any trend, the model fits better if we restrict the period, skipping the period where the pollen concentration is zero or very low. For Urticaceae, and for the sites considered in this paper, this is not necessary in most cases, but it should be done for data from other localities with zero or near zero levels (Belmonte et al.,

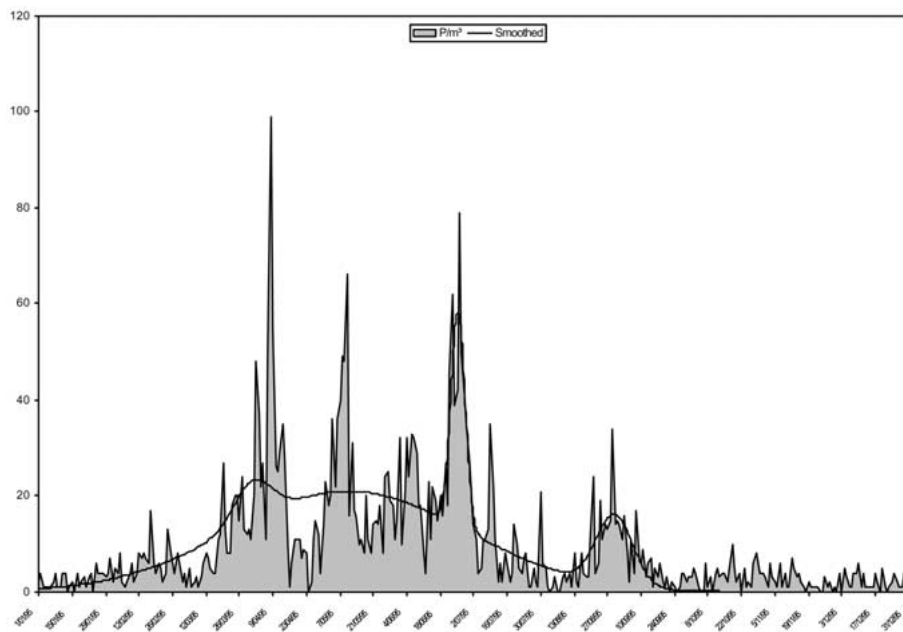


Figure 1. Urticaceae pollen concentrations with trend (Gaussian mixture model), Barcelona, 1995.

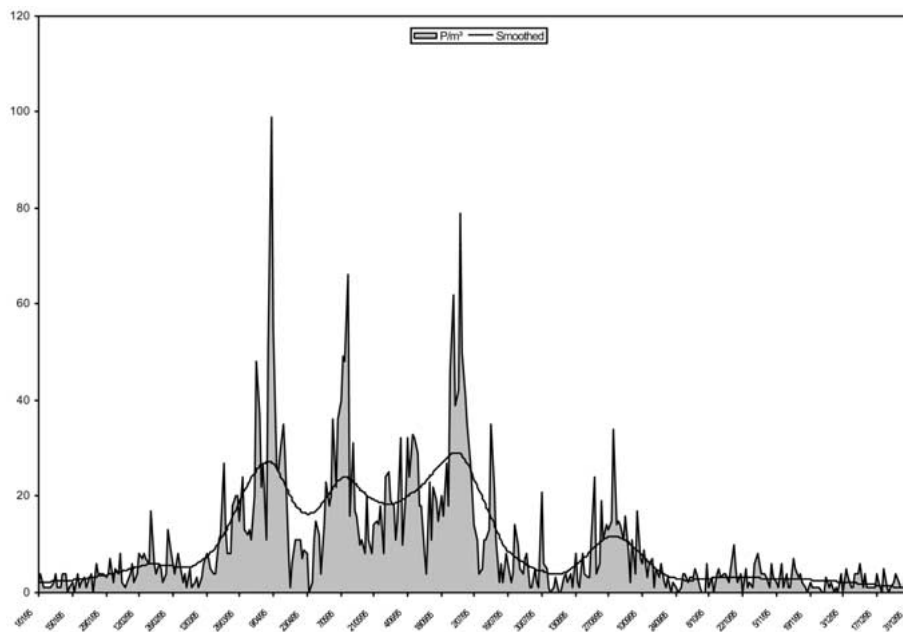


Figure 2. Urticaceae pollen concentrations with trend (Friedman supersmoother), Barcelona, 1995.

1999) and for other pollen types not covered in this paper (e.g. Cupressaceae).

There are two main classes of trends, non-parametric and parametric. For a non-parametric trend, there is no mathematical equation giving $C(t)$ as a function of t . As the rigidity of a mathematical function is avoided, sometimes non-parametric trends can be used without minding the period of time covered by the data. For Urticaceae, in most cases the annual series can be merged into a single series, and one does not need to analyse them separately. The trends obtained are roughly the same. The most used non-parametric smoothing technique is the moving average. For pollen data, a moving average still has a great number of peaks, i.e. it is not smooth enough. More satisfactory results, but still producing a jigsaw curve, are given by other linear filters, like the exponential smoother. We have tried various non-parametric smoothing methods that can be found in the statistical literature (Effromovich, 1999). The best performer, for the pollen series of this paper, is the Friedman's supersmoother (Friedman, 1984), available in the statistical package S-plus.

For parametric trends, there is an equation relating $T(t)$ to t . The equation has some parameters which have to be estimated for each particular series. Most parametric trends used for pollen data are polynomials, e.g. 8th degree polynomials (Comtois and Sherknies, 1991). In this case, t is restricted to the main pollination period or main pollen season period. If not, a polynomial model is not flexible enough to allow for the presence of several zero or almost zero values at the beginning and the end of the year, and produces negative values of $T(t)$.

The approach in this paper is completely different, using a Gaussian mixture model

$$T(t) = K[p_1T_1(t) + p_2T_2(t) + p_3T_3(t) + p_4T_4(t)].$$

Here, K is a scale parameter, which depends on the total annual pollen, the p_i 's are weights (sum equal to 1), and the $T_i(t)$'s are Gaussian density functions,

$$T(t) = (2\pi\sigma^2)^{-1/2}\exp[-(t - \mu)^2/(2\sigma^2)].$$

We adopted a 4-component model because, for Urticaceae pollen, visual inspection of the line plots of the annual series suggests a 4-peaked pattern. This is confirmed by the shape of the non-parametric trends fitted to the series (see Figures 1 and 2). Other pollen types are found only in a part of the year, so that the model would be simpler. Polynomials give

curves that are too simple, losing the 4-peaked pattern, characteristic of Urticaceae pollen in the sites studied. Parameter estimates can be obtained by means of a nonlinear least squares algorithm, but initial estimates are needed. Maximum likelihood methods, like the EM algorithm (McLachlan and Krishnan, 1997) can also be used. We used least squares, choosing initial estimates by visual inspection. Due to the similarity of the series studied in this paper, the same values can be used for different series.

To characterise the residual series $E(t)$, we use ARIMA (Box-Jenkins) models, which can be found in any time series analysis textbook (e.g. Cryer, 1986). Legendre and Legendre (1998) give an introduction aimed to biologists. The auto-regressive (AR) model of order p is

$$E(t) = \phi_1E(t-1) + \dots + \phi_pE(t-p) + W(t),$$

where the ϕ_i 's are the parameters of the model, and $W(t)$ is a purely random, or white noise series (no autocorrelation). All the residual series of this study can be described by an AR model of order 1 or 2, with significant coefficients. More complicated models have been discarded, because they do not lead to a real improvement, and the additional terms are not significant. The selection of the model for each series is based in the autocorrelation structure and the AIC criterion, (Cryer, 1986). The autocorrelation of lag k (ρ_k) is the (linear) correlation of $E(t)$ and $E(t-k)$.

Results and discussion

Summary statistics for Urticaceae pollen daily concentrations are presented in Table 1. The analysis of the relationship between these statistics, reveals a high correlation between mean and standard deviation (0.98), and between mean and maximum (0.93). Similar relationships have been found in other pollen series, e.g. in *Platanus* (Gabarra et al., submitted to Aerobiologia) and Cupressaceae (unpublished). The relationship between the mean and the standard deviation leads to uniformity in the coefficients of variation (SD/mean). To illustrate the relationship between the mean and the maximum, the maximum can be divided by the annual index (sum of mean daily concentrations along the year). This ratio is frequently presented as a percentage in the aerobiological literature. In the series discussed in this paper, the maximum daily concentration falls, in most cases, between 2 and 3% of

Table 1. Aerobiological sampling stations and main geographical and climatic characteristics.

Aerobiological sampling stations (study period)	Geographical characteristics		Climatic characteristics		
	Altitude (m.a.s.l.)	Geographical coordinates	Mean annual temperature (°C)	Annual rain-fall (mm)	Phytoclimates (Allue, 1990)
Barcelona (1994–2000)	12	41°24' N, 02°11' E	16.4	593	Fresh-Tethyc-semiarid
Bellaterra (1994–2000)	190	41°33' N, 02°07' E	15.2	594	Fresh-Continental Oriental-semihumid
Girona (1996–2000)	70	41°59' N, 02°60' E	15.0	740	Fresh-Continental Oriental-semihumid
Lleida (1996–2000)	221	41°37' N, 00°37' E	15.1	385	Fresh-Transitional-semiarid
Manresa (1996–2000)	238	41°43' N, 01°50' E	13.6	619	Fresh-Continental Oriental-semihumid
Tarragona (1996–2000)	20	41°07' N, 01°15' E	15.8	478	Fresh-Tethyc-semiarid

Table 2. Summary statistics for the six sampling sites.

Site	Statistics	1994	1995	1996	1997	1998	1999	2000
Barcelona	No. days	308	365	363	351	365	358	366
	Mean	13.33	9.92	14.27	19.03	10.98	7.95	7.21
	Std. deviation	17.22	12.76	18.41	26.35	14.98	11.98	9.58
	CV	1.29	1.29	1.29	1.38	1.36	1.51	1.33
	Maximum	107.1	99.4	109.2	144.2	106.4	101.5	62.3
Bellaterra	No. days	356	350	345	365	363	358	366
	Mean	9.21	7.47	8.33	13.35	7.13	5.96	4.76
	Std. deviation	11.53	8.85	11.19	17.26	9.19	8.50	7.52
	CV	1.25	1.18	1.34	1.29	1.29	1.43	1.58
	Maximum	56.0	58.1	85.4	84.0	51.1	49.7	45.5
Girona	No. days			306	338	319	352	366
	Mean			22.15	15.16	11.35	15.73	8.31
	Std. deviation			26.98	16.59	14.90	21.49	13.01
	CV			1.22	1.09	1.31	1.37	1.57
	Maximum			138.6	93.6	74.5	112.0	63.7
Lleida	No. days			298	342	359	358	366
	Mean			3.20	2.30	4.15	2.12	1.93
	Std. deviation			3.91	3.39	5.59	3.30	2.77
	CV			1.22	1.47	1.35	1.56	1.43
	Maximum			25.0	18.1	32.2	28.0	16.1
Manresa	No. days			322	353	361	362	366
	Mean			18.85	29.87	13.22	6.32	4.48
	Std. deviation			21.83	38.46	19.30	8.16	5.95
	CV			1.16	1.29	1.46	1.29	1.33
	Maximum			140.7	273.7	149.8	49.0	44.1
Tarragona	No. days			328	361	364	364	359
	Mean			22.09	18.03	14.99	10.0	45.32
	Std. deviation			25.14	22.13	20.30	14.47	8.38
	CV			1.14	1.23	1.35	1.44	1.58
	Maximum			153.3	134.4	131.6	102.9	77.0

Table 3. Spearman correlation values for mean daily Urticaceae pollen concentrations.

	Period	Bellaterra	Girona	Lleida	Manresa	Tarragona
Barcelona	1994	0.780				
	1995	0.644				
	1996	0.761	0.801	0.604	0.556	0.739
	1997	0.867	0.718	0.487	0.798	0.737
	1998	0.774	0.712	0.620	0.749	0.753
	1999	0.786	0.661	0.524	0.613	0.702
	2000	0.699	0.496	0.618	0.655	0.649
Bellaterra	1996		0.759	0.642	0.510	0.651
	1997		0.768	0.518	0.812	0.737
	1998		0.793	0.719	0.759	0.703
	1999		0.720	0.703	0.729	0.642
	2000		0.488	0.574	0.607	0.529
Girona	1996			0.679	0.571	0.683
	1997			0.411	0.857	0.608
	1998			0.795	0.818	0.663
	1999			0.545	0.526	0.656
	2000			0.438	0.509	0.333
Lleida	1996				0.253	0.565
	1997				0.476	0.519
	1998				0.710	0.559
	1999				0.659	0.420
	2000				0.572	0.419
Manresa	1996					0.398
	1997					0.688
	1998					0.722
	1999					0.465
	2000					0.517

the annual total. The least squares estimate for this percentage is 2.32%, with $R^2 = 0.972$. Values below 2% can be attributed to missing data (this could be partially corrected by filling the gaps in the series by means of an imputation algorithm) or to the fact that sometimes the peak concentration is very far from the rest. Values above 3% can be attributed to the existence of several concentrations near the maximum. This is not unusual in Urticaceae pollen, at least at the sites involved in this paper.

Spearman correlation values between series from different sites, for the same year, are presented in Table 3. High correlation between two sites can be taken as an indication of the coincidence of the environmental factors that affect Urticaceae pollen emission and circulation. Thus, the proximity between Barcelona and Bellaterra leads to high, consistent correlation (around 0.7). Correlation between any of these two

sites and Girona is also high, except for year 2000. For other pairs of sites, the correlation is more variable, suggesting occasional coincidence of environmental factors.

In Table 4, we present the parameter estimates for the Gaussian mixture model, in two sites. The K parameters are roughly proportional to means. The other parameters, which are related to the distribution of the pollen along the year, are similar for these two sites, as could be expected from the correlation values of Table 3. Looking at the whole period 1994–2000, the annual distribution is quite stable, in spite of the changes in the annual total. Nevertheless, a somewhat different picture emerges from the results of year 2000, possibly due to the decay in total pollen.

In Figure 1, we present a series $C(t)$, together with the parametric trend $T(t)$. The 4-peaked pattern of these series is evident in this graphical represen-

Table 4. Trend parameter estimates for Barcelona and Bellaterra.

Parameter	1994	1995	1996	1997	1998	1999	2000
Barcelona							
K	4335.20	3358.43	4414.07	6357.60	4223.66	2929.62	2473.68
ρ_1	0.0868	0.0478	0.0005	0.0900	0.068	0.0956	0.1671
μ_1	84.59	90.05	83.86	92.75	92.97	94.56	77.25
σ_1	2.02	7.78	0.70	5.79	1.33	1.80	19.46
ρ_2	0.8686	0.7703	0.8907	0.6517	0.6688	0.7665	0.3989
μ_2	128.64	130.66	143.99	118.59	119.65	124.05	127.26
σ_2	51.55	49.04	35.29	39.20	39.46	48.19	18.70
ρ_3	0.0255	0.0997	0.0464	0.2123	0.0203	0.0627	0.2384
μ_3	152.51	176.45	197.59	193.84	173.52	170.57	166.99
σ_3	1.70	2.98	4.93	6.09	2.25	2.58	9.84
ρ_4	0.0209	0.0822	0.0624	0.0460	0.2429	0.0752	0.1956
μ_4	243.53	241.84	231.14	210.71	263.41	274.09	213.38
σ_4	2.61	7.56	5.10	2.12	117.19	45.36	17.03
Bellaterra							
K	2991.25	2510.12	2981.92	4885.53	2470.07	2099.04	1629.21
ρ_1	0.0756	0.0403	0.0820	0.0372	0.0183	0.0465	0.1020
μ_1	85.02	97.66	105.07	96.75	89.97	95.82	74.51
σ_1	4.27	2.78	7.29	4.28	0.57	2.98	11.60
ρ_2	0.8900	0.8484	0.6783	0.6954	0.8449	0.8553	0.2455
μ_2	135.30	146.31	155.90	116.32	130.52	137.15	121.54
σ_2	40.49	47.77	39.40	39.91	41.52	39.95	11.26
ρ_3	0.0132	0.0400	0.0457	0.1900	0.0351	0.0517	0.4929
μ_3	186.53	175.47	167.30	195.99	173.56	170.46	162.67
σ_3	0.43	0.87	0.84	8.39	4.18	2.47	18.14
ρ_4	0.0212	0.0713	0.1994	0.0774	0.1017	0.0465	0.1596
μ_4	227.04	241.55	271.78	239.47	241.00	242.03	211.15
σ_4	2.79	8.76	39.65	32.25	17.05	19.12	12.60

tation. Figure 2 is the same as Figure 1, but $T(t)$ is the non-parametric trend given by the Friedman super-smoother. Spearman correlation values between trend series from two sites, for the same year, are presented in Table 5. As the trend is assumed to reflect the long term patterns due to environmental factors, these values are higher than those obtained for raw data (Table 3). The non-parametric trend seems to perform better than the Gaussian mixture model, producing correlation values above 0.9. As could be expected, the similarity of the parameters of Table 4 leads to higher values (e.g. in 1997) in Table 5, and conversely (e.g. in 1998).

The residuals $E(t)$ have been calculated by means of the formula $E(t) = X(t)/T(t) - 1$. To shorten the presentation, we have restricted the discussion to

residuals from the non-parametric trend. Of course, there is a danger of $T(t)$ being zero or very close to zero, specially for sites where there is less Urticaceae pollen. In these cases, the series must be trimmed. Nevertheless, no more than 15 days are lost in any case. Spearman correlation values between residual series are also presented in Table 5. For this pair of sites, all correlation values are significant at level $P < 0.01$. For other pairs the correlation is lower, as could be expected from Table 3.

In Table 6, we present the autocorrelations ρ_1 and ρ_2 , and the AR parameters for each residual series. ρ_1 is always positive and significant, while ρ_2 is sometimes helpful to identify the model. A remarkable fact is the negative parameter in the models of year 2000, which do not appear in the other series not reported

Table 5. Spearman correlation values between Barcelona and Bellaterra annual series.

Series	1994	1995	1996	1997	1998	1999	2000
Parametric trend	0.897	0.900	0.804	0.951	0.776	0.863	0.952
Non-parametric trend	0.915	0.952	0.987	0.953	0.914	0.940	0.915
Residuals	0.316	0.270	0.302	0.311	0.181	0.352	0.273

Table 6. First order autocorrelation and AR models fitted to residual series.

Parameter	1994	1995	1996	1997	1998	1999	2000
Barcelona							
ρ_1	0.179	0.148	0.196	0.191	0.274	0.394	0.159
ρ_2	0.094	0.133	0.075	0.085	0.113	9.224	-0.006
ϕ_1	0.184	0.151	0.198	0.197	0.280	0.406	0.255
ϕ_2		0.130					-0.138
Bellaterra							
ρ_1	0.202	0.118	0.191	0.208	0.215	0.246	0.240
ρ_2	0.018	-0.039	0.024	0.136	0.109	0.088	-0.074
ϕ_1	0.223	0.126	0.202	0.189	0.220	0.250	0.277
ϕ_2				0.098			-0.139

in the table. This could be related to lower Spearman correlation values between Barcelona and Bellaterra and the rest of sites for this year.

Conclusions

We can use Spearman correlation as a measure of the similarity of series of different sites in the same year, not only for mean daily pollen concentrations, but also for trends and residuals. High Spearman correlation values between different sites points to analogies in those patterns of climate that affect pollen content in the air. High correlation between trends can be attributed to coincidence in factors that affect to long-term patterns.

Our results lead to a model in which the mean daily pollen concentration can be expressed as the product of some trend by a deviation factor $1 + E(t)$. It is not possible to have a trend with a simple formula, but $E(t)$ is a zero-mean stationary series, which can be characterised in a simple way, using conventional autoregressive time series models. Models for different sites and years can be thus compared. For instance, for Barcelona, in 1994, the model is $E(t) = 0.184 C(t-1) + W(t)$, while for Bellaterra, in the same year, is $E(t) = 0.202 C(t-1) + W(t)$, where $W(t)$ is a purely random

series. These models reflect the short-term behaviour of the pollen concentration, while the trend depends mainly on the climate patterns. The autocorrelation of lag 1 (ρ_1) is positive and significant in all the residual series.

Acknowledgements

The Aerobiological Network of Catalonia (Xarxa Aerobiològica de Catalunya, XAC) is possible thanks to the support of Laboratorios CBF-LETI, S.A.

The authors wish to thank the CICYT for providing financial support for project AMB-0457-CO7-021.

References

- Allue J.L.: 1990, *Atlas fitoclimático de España*. INIA, Madrid.
- Belmonte J., Canela M., Guardans R. and Roure J.M.: 1988, Comparison of pollen data obtained by Cour and modified-Durham methods. *Pollen et Spores* **30**, 257–264.
- Belmonte J., Canela M., Guàrdia R., Guàrdia R.A., Sbai L., Vendrell M., Alba F., Alcázar P., Cabezo B., Gutiérrez M., Méndez J. and Valencia R.: 1999, Aerobiological dynamics of the Urticaceae pollen in Spain, 1992–1998. *Polen* **10**, 10–38.
- Bolós O. de and Vigo J.: 1989, *Flora dels Països Catalans*, Vol. 1. Editorial Barcino, Barcelona, pp. 539–545.

- Bringfelt, B., Engström, I and Nilsson, S: 1982, An evaluation of of some models to predict airborne pollen concentration from meteorological conditions in Stockholm, Sweden. *Grana* **21**, 59–64.
- Comtois P. and Sherknies D.: 1991, Pollen curves typology. *Grana* **30**, 184–189.
- Comtois P. and Sherknies D.: 1992, Le pollen de l'Ambroisie (*Ambrosia artemisiifolia* L.): Previsions et prévention. *Allergie et Immunologie* **24**, 22–24.
- Cryer J.D.: 1986, *Time Series Analysis*. PWS-Kent Pub. Co, Boston.
- Domínguez E., Galán C., Villamandos F. and Infante F.: 1991, Manejo y evaluación de los datos obtenidos en los muestreos aerobiológicos. *Monografías REA/EAN* **1**, 11–18.
- Effromovich S.: 1999, *Nonparametric Curve Estimation*. Springer: New York.
- Friedman J.H.: 1984, A variable span smoother. *Technical Report* 5. Laboratory for Computational Statistics, Dept. of Statistics, Stanford Univ., Stanford.
- Hirst J.M.: 1952, An automatic volumetric spore-trap. *Ann Appl Biol* **36**, 257–265.
- Legendre L. and Legendre P.: 1998, *Numerical Ecology*. Elsevier: Amsterdam.
- McLachlan G.T. and Krishnan T.: 1997, *The EM Algorithm and Extensions*. Wiley: New York.
- Moseholm L., Weeke E.R. and Petersen B.N.: 1987, Forecast of pollen concentrations of Poaceae (grasses) in the air by time series analysis. *Pollen et Spores* **29**, 305–322.
- Starck, P.C., Ryan, L.M., McDonald, J.L. and Burge, H.A.: 1997, Using meteorological data to predict daily ragweed pollen levels. *Aerobiologia* **13**, 177–184.

