

## Detection of chaos: new approach to atmospheric pollen time-series analysis

M.M. Bianchi<sup>1</sup>, C.M. Arizmendi<sup>2</sup>, and J.R. Sanchez<sup>2</sup>

<sup>1</sup> Laboratorio de Palinología, Departamento Biología, Facultad de Ciencias Exactas y Naturales, Universidad Nacional de Mar del Plata, 7600 Mar del Plata, Argentina

<sup>2</sup> Departamento de Física, Facultad de Ingeniería

Received January 25, 1991; revised February 24, 1992; accepted February 26, 1992

**Abstract.** Pollen and spores are biological particles that are ubiquitous to the atmosphere and are pathologically significant, causing plant diseases and inhalant allergies. One of the main objectives of aerobiological surveys is forecasting. Prediction models are required in order to apply aerobiological knowledge to medical or agricultural practice; a necessary condition of these models is not to be chaotic. The existence of chaos is detected through the analysis of a time series. The time series comprises hourly counts of atmospheric pollen grains obtained using a Burkard spore trap from 1987 to 1989 at Mar del Plata. Abraham's method to obtain the correlation dimension was applied. A low and fractal dimension shows chaotic dynamics. The predictability of models for atmospheric pollen forecasting is discussed.

**Key words:** Airborne pollen forecasting – Chaos – Time series

### Introduction

Pollen and spores are the most abundant biological particles in the atmosphere, with concentrations near the ground ranging from a few hundreds to several thousands of particles  $m^{-3}$  (Comtois and Sherknies 1987). Pollination is both a natural phenomenon and one of the processes responsible for hay fever; hence, many researchers have tried to apply forecasting algorithms to sequential pollen counts in order to predict future airborne concentrations. In spite of this, forecasts have been persistently shown to be less than 60% accurate (Comtois et al. 1989).

Models to forecast pollen concentration in the air are principally based on pollen and atmospheric weather interactions. Several statistical techniques (O'Rourke 1986, 1989; Goldberg et al. 1988; Comtois and Sherknies 1987; Comtois et al. 1989; Leuschner 1990) have been

used to predict atmospheric pollen concentrations from prevailing weather conditions and from the weather conditions of recent previous days. In this regard, the most successful attempts have taken into account the time-series characteristic of pollen data (Comtois and Sherknies 1987). Attempts to predict pollen concentration in the air by means of phenological data have also been made (Driessen et al. 1989). On the other hand, to our knowledge not much work on deterministic models (Moseholm et al. 1987; Goldberg et al. 1988) has been done.

Although a deterministic model would be useful, a previous question that must be answered is whether the intrinsic nature of the atmospheric pollen record allows prediction modelling. Recent detection of chaos in several natural systems (May 1976; Procaccia 1988) has shown that prediction in nature is not always possible even in systems described by deterministic models with a few degrees of freedom. As dynamic properties of atmospheric pollen systems are still unknown, a proper answer to the above question will therefore be given by detection of chaos in the atmospheric system. Detection of deterministic chaos is given by the correlation dimension introduced by Grassberger and Procaccia (1983).

### Materials and methods

The material used in this study was obtained with an automatic and volumetric Burkard pollen and spore trap, situated at the roof of the Facultad de Ciencias Exactas y Naturales, 12 m above ground level. The area surrounding the sample is typical of Mar del Plata.

The rate of air flow of 10 l/min is obtained by sucking air through a  $14 \times 2 \text{ mm}^2$  orifice, always orientated against the wind flow; the sucking rate is checked weekly. Behind the slit, a drum rotates at a speed of 2 mm/h. The particles are collected on a cellophane tape (Melinex), which is 19 mm wide and located just below the orifice. The sticky collecting surface comprises nine parts vaseline: one part paraffin in toluene. The exposed tape is removed from the drum, cut into pieces of 48 mm corresponding to 24-h intervals, then embedded into a solution of polyvinylalcohol (Gelvatol), water and glycerol and covered with a cover glass. Slides were studied as 12 transects per day. The pollen was counted at

**Table 1.** Statistics of atmospheric pollen hourly counts for the first sampling cycle (August 1987–August 1988; number of observations = 354)

| Hours | Mean  | SD    | Min | Max |
|-------|-------|-------|-----|-----|
| H2    | 3.80  | 7.92  | 0   | 64  |
| H4    | 4.53  | 11.26 | 0   | 107 |
| H6    | 5.88  | 19.43 | 0   | 251 |
| H8    | 7.16  | 21.11 | 0   | 230 |
| H10   | 9.41  | 31.52 | 0   | 318 |
| H12   | 10.38 | 31.70 | 0   | 363 |
| H14   | 9.42  | 28.54 | 0   | 299 |
| H16   | 5.73  | 11.25 | 0   | 115 |
| H18   | 4.29  | 8.98  | 0   | 79  |
| H20   | 3.34  | 6.55  | 0   | 72  |
| H22   | 3.21  | 6.16  | 0   | 58  |
| H24   | 3.43  | 8.36  | 0   | 101 |

**Table 2.** Statistics of atmospheric pollen hourly counts for the second sampling cycle (August 1988–August 1989; number of observations = 309)

| Hours | Mean  | SD    | Min | Max |
|-------|-------|-------|-----|-----|
| H2    | 7.99  | 46.64 | 0   | 760 |
| H4    | 8.52  | 38.82 | 0   | 573 |
| H6    | 10.03 | 35.44 | 0   | 410 |
| H8    | 12.49 | 42.31 | 0   | 371 |
| H10   | 11.31 | 35.55 | 0   | 512 |
| H12   | 11.94 | 44.98 | 0   | 647 |
| H14   | 11.06 | 41.94 | 0   | 558 |
| H16   | 6.23  | 19.69 | 0   | 296 |
| H18   | 4.21  | 11.99 | 0   | 168 |
| H20   | 3.26  | 7.01  | 0   | 59  |
| H22   | 3.84  | 8.49  | 0   | 75  |
| H24   | 5.19  | 20.64 | 0   | 215 |

a magnification of  $\times 400$  for the first year cycle (August 1987–8) and at  $\times 200$  for the second (August 1988–9), and corresponding to 13.5 and 27 min of sampling every 2 h respectively. The method of counting pollen follows that of Kapyla and Penttinen (1981). Hourly counts were stored in a database file for further analysis. Statistics of hourly counts of the first year are shown in Table 1 and those for the second year in Table 2.

## Chaos

Since Poincare (1892), the often repeated statement that given the initial conditions we know what a deterministic system will do in the future, is well known to be not always true. Even with the simplest equations of motion, any nonlinear system with chaotic behaviour with infinitesimally different starting points, will end up with *wildly* different outcomes. Traditionally, systems with a finite number of degrees of freedom have been divided into those with few degrees of freedom, which were supposed to exhibit some kind of regular ordered motions, and those which are unpredictable because of the large number of degrees of freedom. The dynamics of the latter ones must be treated statistically.

The past few decades have seen a complete change of view, which affects almost all the practical applica-

tions. The dynamics of a system are usually neither completely regular nor properly described by the methods of statistics. A typical system is mixed: it exhibits regular or chaotic motion for different initial conditions and the transition between the two types of motions is complicated, subtle and beautiful.

The simple models, commonly known as dynamical systems, which present this behaviour are ordinary differential equations like

$$dx/dt = F_{\mu}(x) \quad (1)$$

where  $x$  is a point in a  $d$ -dimensional space of time-dependent dynamic variables and  $\mu$  is a point in the space of time-independent parameters, typically controlled by the experimenter.  $F_{\mu}(x)$  is some nonlinear function of the dynamic variables. The dimension of the dynamic system is the dimension of the space of variables. The solutions of Eq. 1 may be stationary, periodic or *chaotic*, depending on the region of the parameter space where  $\mu$  is. The way that dynamical systems become chaotic by following a path in the parameter space is the same for completely different natural systems. This fact, known as *universality*, was demonstrated by Feigenbaum (1978) for the logistic map

$$x_{n+1} = \mu x_n (1 - x_n), \quad (2)$$

commonly used as a model of population growth for a seasonally breeding population in which generations do not overlap (May 1976). The logistic map relates the population in generation  $n+1$  with the population in the previous generation  $n$ . The route to chaos followed by the variation of the parameter  $\mu$  is characterised by a cascade of period doublings. When the parameter  $\mu$  is small the logistic map has a unique fixed point that bifurcates in 2, 4, 8, ...,  $2^n$ , as  $\mu$  increases. Feigenbaum found that the ratio between the differences of the successive bifurcation parameter values and the ratio between the solution splitting are universal (i.e. independent of the particular system). Chaos is reached after the bifurcation scheme ( $\mu = 3.57$ ) (May 1976).

This is the most characteristic route to chaos. It has been found in hydrodynamic experiments associated with the onset of turbulence (Libchaber and Maurer 1981), electronic circuits with nonlinear devices as diodes or transistors (Testa et al. 1982), the Belousov-Zhabotinsky chemical reaction (Simoyi et al. 1982), ecological models like Lotka-Volterra equations, etc. A few months ago, Tilman and Wedin (1991) proposed a model for the dynamics of a perennial grass which presents this route to chaos.

To understand the nature of the dynamic complexity involved in so-called chaotic motion, consider the logistic map (2) with  $\mu = 4$ . The behaviour here is known as *full chaos*, because the chaotic region extends to the whole interval  $[0, 1]$ . Take the initial conditions such that  $0 < x_0 < 1$ . Every initial condition  $x_0$  gives rise to an orbit  $x_0, x_1 = f(x_0), x_2 = f[f(x_0)] \dots$ . In this sense this dynamic system is deterministic and the future evolution may be known in principle. However if you pick any nearby initial conditions  $x_0$  and  $x'_0$ , such that  $|x_0 - x'_0| = \epsilon$

is very small, and iterate (1), on iteration  $n$ , the distance between the orbits increases as  $|x_n - x'_n| = 2^n \varepsilon$ , i.e. it grows exponentially. Thus, predictability does not exist in practice. This is known as sensitivity to initial conditions and is characteristic of chaos.

Chaotic motion appears in dynamical systems when local exponential divergence of trajectories is accompanied by global confinement in the phase space but, because of the confinement, this stretching cannot continue without folding. Repeated folding and refolding produces very complicated behaviour which is described as chaotic. In multidimensional dissipative systems, phase space volume shrinks in some directions and grows and folds onto itself in others so the attractor, while thin, can be very complicated. The attractor is the region of the phase space to which the trajectories are confined at asymptotic times.

### Dimensions of strange attractors

The attractor can be a point (with 0 dimension, e.g. the pendulum with friction), a curve (with dimension 1, e.g. an oscillator), or a *strange* attractor characterized by *noninteger* dimension (Ott 1981). The relevant definition of dimension is that due to Hausdorff (see for example Mandelbrot 1977)

$$d = \lim_{\varepsilon \rightarrow 0} \ln N(\varepsilon) / \ln(1/\varepsilon), \quad (3)$$

where, if the attractor is a subset of a  $p$ -dimensional ordinary space,  $N(\varepsilon)$  is the number of  $p$ -dimensional cubes of side  $\varepsilon$  needed to cover the attractor.

As an example of the application of Eq. 3, if the attractor is a fixed point, then  $N(\varepsilon) = 1$  and according to Eq. 3, the Hausdorff dimension is 0. If the attractor is a segment of unitary length then  $N(\varepsilon) = 1/\varepsilon$  and the dimension is 1. Dissipative dynamical systems which exhibit chaotic behaviour have an attractor in phase space which is *strange*. Strange attractors are typically characterized by *fractal* dimensionality, which is smaller than the number of degrees of freedom. A correlation dimension related to the fractal dimension can also be defined on a strange attractor.

The correlation dimension has been introduced by Grassberger and Procaccia in a remarkable paper (Grassberger and Procaccia 1983), with an algorithm to obtain it from the time series of a single observable. Since then it has been largely used because of its reliability and simplicity. The measure is obtained by considering correlations between points of a long time series on the attractor. Denote the  $N$  points of such series by

$$\{\vec{x}_i\}_{i=1}^N \equiv \{\vec{x}_{t+i\tau}\}_{i=1}^N \quad (4)$$

where  $\tau$  is an arbitrary but fixed time increment.

The definition of the correlation integral is

$$C(r) = \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{i,j=1}^N \theta(r - |\vec{x}_i - \vec{x}_j|) = \int_0^r d^d r' c(r') \quad (5)$$

where  $\theta(x)$  is the Heaviside function and  $c(r)$  is the standard correlation function.  $C(r)$  behaves as a power of  $r$  for small  $r$ :

$$C(r) \sim r^v \quad (6)$$

where the exponent  $v$  is the correlation dimension as defined by Grassberger and Procaccia (1983). Improvements to this technique of calculation have been supplied by Abraham et al. (1986), reducing the number of members of the time series.

### Correlation dimension of the pollen time-series attractor

The algorithm of Abraham et al. (1986) was applied to the pollen time series. Equation 6 indicates that the correlation dimension can be obtained as the slope of log

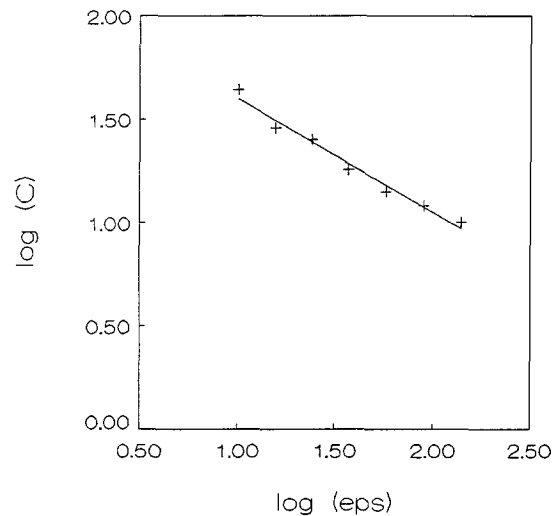


Fig. 1. Graph of  $\log(C)$  vs  $\log(\varepsilon)$  for the first year's observations. The line through the points has a slope = 0.55. The slope is the correlation dimension

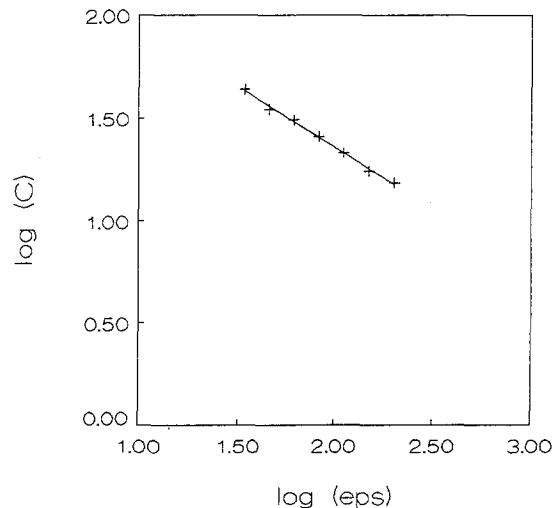


Fig. 2. Graph of  $\log(C)$  vs  $\log(\varepsilon)$  for the second year's observations. The line through the points has a slope = 0.59

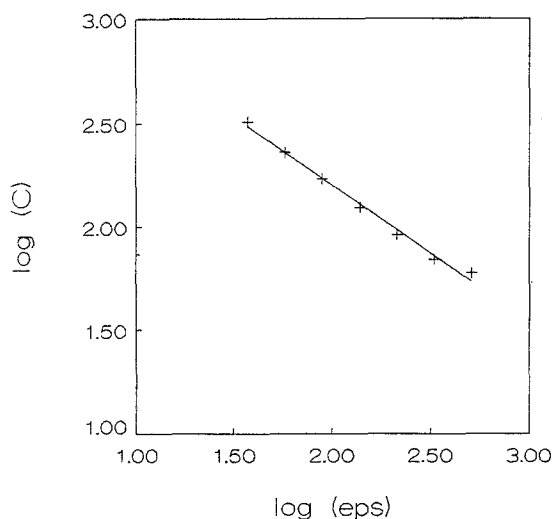


**Fig. 3.** Graph of  $\log(C)$  vs  $\log(\epsilon)$  for two consecutive year's observations. The line through the points has a slope = 0.66

( $C$ ) vs  $\log(\epsilon)$ . As the pollen time series extends for 2 years, we calculated the correlation dimension for the time series corresponding to the first and second year samplings separately. The results of the first year are shown in Fig. 1; those for the next year are shown in Fig. 2. The slope of the curve for both figures is the same. To confirm this result we also calculated the correlation dimension for the complete time series for both years. The curve obtained is displayed in Fig. 3 and its slope coincides with the previous one. The correlation dimension corresponding to the attractor of the pollen time series obtained with this method is  $v = 0.66 \pm 0.01$ . A canonical example of the chaotic map with correlation dimension less than one is the logistic map (Grassberger and Procaccia 1983; Abraham et al. 1986).

## Discussion

A new analysis of the dynamic characteristic of atmospheric pollen time series obtained by a standard method of aerobiology was carried. For this purpose, a strange attractor was sought of this time series through the study of the correlation dimension of Grassberger and Procaccia. The value of the dimension found was non integer and less than one, indicating that the system can be described by a one-dimensional chaotic map. This chaotic behaviour indicates that there are serious restrictions for pollen predictions. These restrictions are of the same kind as those that appear in atmospheric predictions associated with chaos found by Lorenz (1964). This unpredictability, related with the chaotic dynamics and not with a great number of relevant variables in the system, places a fundamental limit on long-term prediction while suggesting possibilities for short-term prediction (Farmer and Sidorowich 1987). Forecasting is a money- and time-consuming activity. In spite of this, the accuracy obtained is far from satisfactory. A previous analysis of

the kind developed in this paper would help when making decisions concerning the investment on forecasting.

**Acknowledgements.** This work was partially supported by CON-ICET, Alexander von Humboldt Foundation and Volkswagenwerk Foundation.

## References

- Abraham NB, Albano AM, Das B, De Guzman G, Young S, Giorggia RS, Puccioni GP, Tredicce JR (1986) Calculating the dimension of attractors from small data sets. *Phys Lett* 114A: 217–221
- Comtois P, Sherknies D (1987) An aerobiological model for pollen forecasting. 18th Conference on Agricultural and Forest Meteorology, 8th Conference on Biometeorology and Aerobiology, W. Lafayette, Indiana. American Meteorological Society, Boston, Mass
- Comtois P, Batchelder G, Sherknies D (1989) Pre-season forecasting. In: Comtois P (ed) *Aerobiology health environment: a symposium*, Université de Montréal
- Driessen MNB, Van Herpen RMA, Moelands RPM, Spijkma FThM (1989) Prediction of the start of grass pollen season for the western part of the Netherlands. *Grana* 28:37–44
- Farmer JD, Sidorowich JJ (1987) Predicting chaotic time series. *Phys Rev Lett* 59:845–848
- Feigenbaum MJ (1978) Quantitative universality for a class of nonlinear transformations. *J Stat Phys* 19:25–52
- Goldberg C, Buch H, Moseholm L, Weeke E (1988) Airborne pollen records in Denmark, 1977–1986. *Grana* 27:209–218
- Grassberger P, Procaccia I (1983) Characterization of strange attractors. *Phys Rev Lett* 50:346–349
- Käpylä M, Penttinen A (1981) An evaluation of microscopical counting methods of the tape in Hurst-Burkard pollen and spore trap. *Grana* 20:131–141
- Leuschner R (1990) Twenty-one year of airborne pollen determinations. *J Palynol* 91:305–320
- Libchaber A, Maurer J (1981) A Rayleigh-Bénard experiment: helium in a small box. In: Riste T (ed) *Nonlinear phenomena at phase transitions and instabilities*. Plenum, New York, p 259
- Lorenz EN (1964) The problem of deducing the climate from the governing equations. *Tellus* 16:1
- Mandelbrot BB (1977) *Fractals*. Freeman, San Francisco
- May R (1976) Simple mathematical models with very complicated dynamics. *Nature* 261:459–467
- Moseholm L, Weeke E, Peterson B (1987) Forecast of pollen concentration of Poaceae (Grasses) in the air by time series analysis. *Pollen et Spores* 29:305–322
- O'Rourke M (1986) The implications of atmospheric pollen rain for pollen profiles in the arid southwest. PhD Thesis, University of Arizona
- O'Rourke M (1989) Relationship between airborne pollen concentration and weather parameters in an arid environment. In: Comtois P (ed) *Aerobiology, health, environment: a symposium*, Université de Montréal
- Ott E (1981) Strange attractors and chaotic motions of dynamical systems. *Rev Mod Phys* 53:655–671
- Poincaré H (1892) *Les méthodes nouvelle de la mécanique céleste*. Gauthier-Villiar, Paris
- Procaccia I (1988) Universal properties of dynamically complex systems: the organization of chaos. *Nature* 333:618–623
- Simoyi RH, Wolf A, Swinney HL (1982) One-dimensional dynamics in a multicomponent chemical reaction. *Phys Rev Lett* 49:245
- Testa JS, Pérez J, Jeffries C (1982) Evidence for universal chaotic behaviour of a driven nonlinear oscillator. *Phys Rev Lett* 48:714
- Tilman D, Wedin D (1991) Oscillations and chaos in the dynamics of perennial grass. *Nature* 353:653–655