STUDY ON THE GEOMETRICAL AND DYNAMICAL CHARACTERISTICS OF THE AROSA OZONE SERIES ATTRACTOR

VASILE CUCULEANU[1], CONSTANTIN RADA[2], ALEXANDRU LUPU[3]

[1]*Physics Faculty, Bucharest University, P.O. Box MG-11, Măgurele, Bucharest, Romania*
[2]*National Meteorological Administration, Şos. Bucureşti-Ploieşti 97, Bucharest, Romania*
[3]*York University, Centre for Research in Earth and Space Sciences, 4700 Keele St., Toronto, Ontario M3J 1P3, Canada*

Étude sur les caractéristiques géométriques et dynamiques de l'attracteur de la série d'ozone totale Arosa. On a étudié la présence du chaos déterministe dans la dynamique de la série des valeurs d'ozone totalement mesuré à la station Arosa en utilisant la théorie des sistèmes dynamiques. En complétant les données qui manquent de cette série par les valeurs moyennes diurnes multiannuelles on a obtenu une série à pas de temps régulé d'un jour dans la période 1926–2005 et on a estimé les caractéristiques géométriques et dynamiques de l'attracteur: la dimension de corrélation, l'exponent Lyapunov et l'entropie. La valeur fractale de la dimension de corrélation et les valeurs positives de l'exponent Lyapunov et de l'entropie constituent des arguments en faveur de la présence du chaos déterministe dans la dynamique de la série Arosa. Dans ce cas, on peut faire la prédiction de type déterministe des valeurs d'ozone total seulement sur un intervalle égal à l'inverse de l'entropie. Outre cet interval on peut faire la prédiction seulement du point de vue statistique.

Key words: total ozon, time series, deterministic chaos, correlation dimension, Lyapunov exponent, entropy.

1. INTRODUCTION

The main objective of this paper is to search for the existence of deterministic chaos in the total ozone dynamics. In order to achieve this goal it is necessary to have an adequate time series of total ozone values which are spaced uniformly in time and contains enough data such as to allow for a reliable estimation of the characteristics describing the attractor dynamics. The total amount of ozone in a column from the surface to the top of the atmosphere, referred to as total ozone, is determined by sun light spectrometry. The Dobson spectrophotometer has been the first instrument used for this purpose. A Dobson unit (1 DU) is defined as being a 0.01 mm layer of pure ozone at standard pressure and temperature. The range of typical values for the atmospheric total ozone is 200−400 DU. The length (1926−2005) and quality of the measurements make the Arosa time series a convenient series for studying the characteristics of the total ozone dynamics in the atmosphere. This series has been downloaded from the web site of the Institute for Atmospheric and Climate Science, Zurich, Switzerland (http://www.iac.ethz.ch/en/research/chemie/tpeter/totozon.html). An important issue of a total ozone series is the fact that the respective records are not continuous in the sense that no regular time interval (one day) exists among all measured values, the percentage of missing data being about 36% (Janosi, Muller, 2005). This is explicable for a spectrometric method relying on the direct sun light observations in the UV range, because clouds hamper the performance of precise measurements.

The analysis of a time series by the dynamical systems concepts requires regular time step among the sampled data, because for the time series consisting in data which are not spaced uniformly in time, the embedding theorems do not apply, and the use of time delay embedding nevertheless does not lead to a correct insight into the data structure. The gap-filling procedures used in this paper are based on the calendar day averages obtained from the total number of the respective day existing in the time series.

Study on the dynamical behavior of total atmospheric ozone measured by means of the Total Ozone Mapping Spectrometer (TOMS) on board the Nimbus-7 satellite suggests evidence for low dimensional ozone attractors, their correlation dimensionalities being between 3 and 7 (Yang *et al.*, 1994).

In Section 2, the methodology to process the Arosa time series is presented. In Section 3, the characteristics indicating the nature of a time series dynamics are introduced. In Sections 4 and 5 the obtained results and corresponding conclusions and discussions are presented.

2. DATA

The Arosa time series cover the time interval from July 1926 to May 2005 (http://www.iac.ethz.ch/en/research/chemie/tpeter/totozon.html). In order to study the dynamics of the entire Arosa time series, appropriate gap filling methods have to be used. Since there are time intervals of missing data of the order of months (29.09.1988–24.11.1988) or years (11.06.1929–19.08.1931) the usual interpolation methods cannot be used. An alternative procedure to fill the time intervals of missing data could be the use of the multiannual averages of the calendar day values of total ozone. This procedure may be justified by the fact that, since the Arosa time series cover a great number of years, from a physical point of view, the respective multiannual averages are quite suitable approximations for the real data which are missing. This procedure has been used to generate a complete Arosa time series for the period 1926−2005 (time series T_1) containing 28,822 data. Taking into account the fact that a period longer than two consecutive years has been filled with the same multiannual averages, in order not to cause a possible alteration of the dynamics of the entire series, a time series (T_2) for the period 1931−2005, with 26,968 data, has been obtained from T_1 by excluding the time interval 1926−1930.

3. CHARACTERISTICS INDICATING CHAOTIC BEHAVIOR

In order to make evident the presence of the chaos in the dynamics of a time series, the correlation dimension, the Lyapunov exponents, and the entropy have to be determined. A qualitative indication upon the dynamics underlying the time series is provided by the power spectrum which has to be broadband in the case of a chaotic systems. On the other hand, the onset of the broadband spectrum cannot always be considered as an argument in favour of the existence of chaos, because noisy periodic and quasiperiodic signals can be characterized by a broadband spectrum, too.

The chaotic dynamics of a time series is evidenced by a broadband power spectrum, a fractal correlation dimension, at least one positive Lyapunov exponent, and a positive finite metric entropy. In this case, a strange attractor underlies the temporal behaviour of the respective series.

3.1. PHASE SPACE RECONSTRUCTION

Let $\{x_i = x(t_i)\}_{i=1,N}$ represent the time series of atmospheric total ozone with all gaps filled, where $t_i = t_{i-1}+\Delta t$, t_1 is the starting time of the measurements, Δt is the sampling time (one day in the case of the Arosa time series) and N is the total number of data.

It has been shown by Takens (Takens, 1981) that from a single time series one can properly reconstruct an m-dimensional phase space by taking the original time series $x(t_i)$ and its successive time shifts (delays) as coordinates of a vector time series given by

$$X_i = \{x(t_i), \mathrm{x}(t_i + \tau), \ldots \mathrm{x}[t_i + (m\text{-}1)\tau]\} \tag{1}$$

where m is the dimension of the vector X_i, often referred to as the embedding dimension, and τ is the time delay. If τ is properly chosen, the variables $x(t_i)$, $\mathrm{x}(t_i + \tau)$, $\ldots\mathrm{x}[t_i + (m\text{-}1)\tau]$ will be independent, this being all one needs to define a phase space. The typical choice of τ is based on the decorrelation time of the time series. This is defined as the lag time at which the autocorrelation function first falls bellow a threshold value which is commonly taken as 1/e in meteorology (Zeng *et al.*, 1992). Other common choices include the first zero and the first inflexion point of the autocorrelation function (Yang *et al.*, 1994).

3.2. CORRELATION DIMENSION

For a fixed embedding dimension m, the time series of vectors X_i in the embedding space are used to calculate the correlation sum, $C_m(r,N)$, defined as the fraction of all possible pairs of points (X_j, X_k) which are closer than a given distance r in a particular norm:

$$C_m(r,N) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\theta(r - \|X_i - X_j\|) \tag{2}$$

where $\theta(a)$ is the Heaviside function which reaches 0 or 1 when $a \leq 0$ or $a > 0$, respectively, r is a given positive number, $\|X_i - X_j\|$ stands for the distance between two points X_i , X_j in the embedding space and N is number of data.

The correlation dimension d_c is defined as the slope of the correlation sum in the log-log coordinate system (Grassberger, Procaccia, 1983):

$$d_c = \lim_{r\to 0}\lim_{N\to\infty} \log C_m(r,N)/\log r \tag{3}$$

Practically, the calculation of the correlation dimensions has to be repeated with increasing m until the estimates of d_c no longer change significantly. In this case, the estimates obtained are the correlation dimension D_c, and the corresponding value of the embedding dimension is called the saturated embedding dimension and is denoted by M.

D_c and M give the upper and lower limits of the number of essential variables necessary to simulate the dynamics of the system (Yang *et al.*, 1994).

3.3. ENTROPY

When the correlation dimension no longer changes, for a sufficiently large m ($m>M$), the system entropy, K, can be expressed as (Zeng, Pielke, 1993):

$$K = \frac{1}{n\tau}\ln\frac{C_m(r)}{C_{m+n}(r)} \tag{4}$$

where r has to be within the linear region of the log $C_m(r)$ versus log r plot for each embedding dimension m. The reciprocal of the system entropy is roughly the time interval over which a deterministic prediction is possible in case of a strange attractor.

3.4. LYAPUNOV EXPONENTS

The Lyapunov exponent is a quantitative measure of the rate at which nearby trajectories in phase space diverge.There exists as many Lyapunov exponents as phase space dimensions. A strange attractor has at least one positive Lypunov exponent. In addition, for any continuous chaotic sysytem, there must be at least one exponent equal to zero. Let us consider the most important exponent, *i.e.*, the maximal Lyapunov exponent .

Let X_{n1} and X_{n2} be two points in phase space with distance $\|X_{n1}-X_{n2}\| = \delta_0 << 1$. Denote by $\delta_{\Delta n}$ the distance at time n between the two trajectories having these points as origin, $\delta_{\Delta n} = \|X_{n1+\Delta n} - X_{n2+\Delta n}\|$. The exponent λ is determined by (Kantz, Schreiber, 1997):

$$\delta_{\Delta n} \cong \delta_0\, e^{\lambda\delta\Delta n},\ \delta_{\Delta n} << 1,\ \Delta n >> 1$$

If λ is positive, the nearby trajectories diverge exponentially, which means that the the time series has a chaotic dynamics. In this paper only the largest positive exponent, λ, has been estimated by making use of the algorithm proposed by Wolf *et al.* (1985). The sum of the positive Lyapunov exponents is an estimation of the entropy.

For most of the calculations presented in this paper, the Chaos Data Analyzer (Sprott, Rowlands, 1995) was used.

4. RESULTS

4.1. AUTOCORRELATION FUNCTION

Figure 1 shows the Arosa time series for the period 1926−2005, with all missing points filled in (time series T_1) and Figure 2 shows the corresponding autocorrelation function. It can be seen that the autocorrelation function exhibits a slowly attenuated oscillation pattern and has a clear annual cycle. This character is common to many weather and climate systems. As was stated in section 3.1, in order to determine the delay time, three procedures based on autocorrelation function can be used: e-fold time (τ_1), the first zero(τ_2) and the first inflexion point(τ_3) of this function. In case of series T_1 these quantities are: $\tau_1 = 56.1$ d, $\tau_2 = 96$ d, $\tau_3 = 1÷2$ d. The autocorrelation function of series T_2 is identical with that of T_1, the only difference being $\tau_1 = 55.5$ d. The fact that the autocorrelation does not fall to zero within a very short time is an indication that the time series does not display the sign of a completely random behaviour, because for Gaussian white noise, the zero of the autocorrelation function is immediately attained.

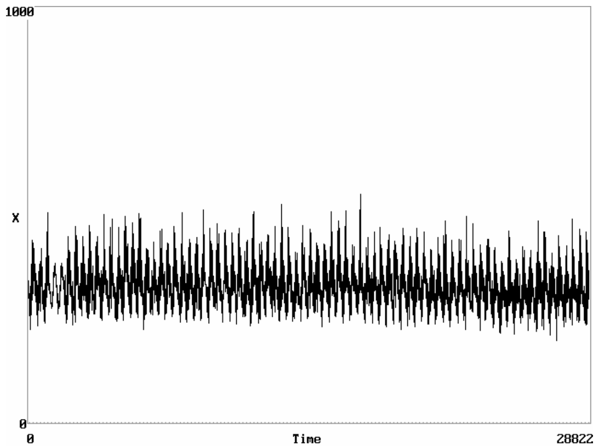

Fig. 1 – The complete Arosa time series for the period 1926–2005.

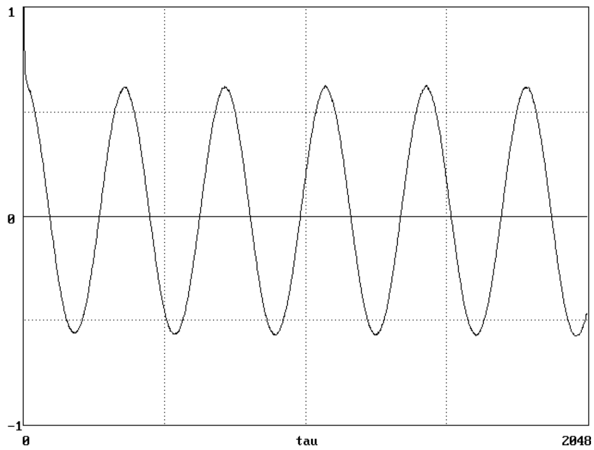

Fig. 2 – Autocorrelation function of the time series T_1 and T_2.

4.2. POWER SPECTRUM

The daily averages of the total ozone column at Arosa show a seasonal variation with larger values in spring and lower in fall and a day-to-day variability. The systematic seasonal variation is due to the general circulation in the stratosphere while the day-to-day variation is related to the meteorological conditions.

The one day-resolution of the total ozone data allows for capture of the dynamics of the atmospheric processes characterized by seasonal variation such as spring maxima and fall minima.The power spectrum has been computed by making use of the fast Fourier transform method. Figure 3 shows the power spectrum of series T_1, which is identical with that of the series T_2. It can be seen that the power spectra are broadband with fluctuations due to noise and display a clear peak value (dominant frequency equals to 0.0039d^{-1}) corresponding to the seasonal variability. Chaotic systems are characterized by a broadband spectrum. But this is only a necessary condition for the existence of chaos, because noisy periodic and quasi-periodic signals can also be characterized by a broadband power spectrum. Therefore, in order to verify the existence of chaos and extract the dynamics from the time series, it is necessary to calculate the correlation dimension and the Lyapunov exponent.

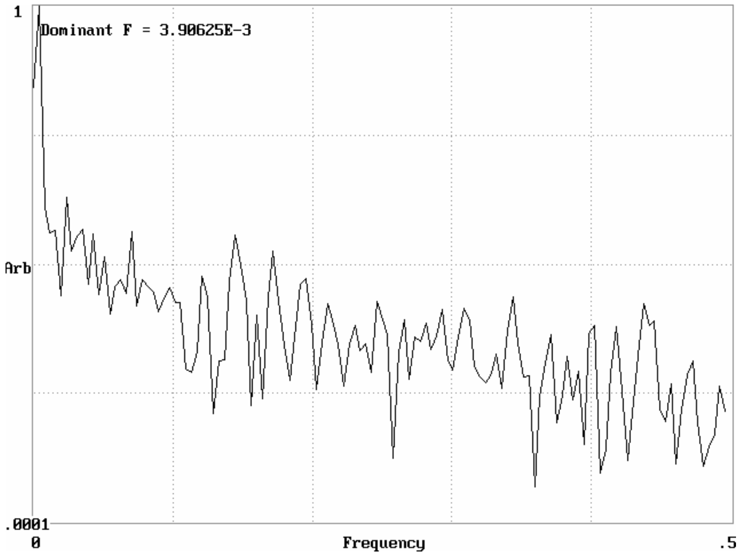

Fig. 3 – Power spectrum for the series T_1 and T_2.

4.3. CORRELATION DIMENSION, ENTROPY AND LYAPUNOV EXPONENTS

A particular problem specific to long time series of physical data, is the possible lack of stationarity. In case of time series T_1, the nonstationarity could be due to the fact that the abundance of ozone has been changing with time, most likely due to external forcing (*e.g.*, release in the atmosphere of chlorofluorocarbons by industry; solar cycle effects). The test for stationarity that we have applied consists in splitting the time series into two halves and verifying that the quantities d_c calculated for the first half agree with those calculated for the second half and with those calculated for the entire time series (Sprott, Rowlands, 1995). By applying this test to series T_1 , significant differences between the values of d_c for the two halves have been pointed out. Thus, before making a calculation of d_c, the time series were detrended by making use of the maximum-entropy method (MEM) (Ghil *et al.*, 2002) and of the differentiation (D) (*i.e.*, the difference of successive data values – Sprott, Rowlands, 1995). Using the MEM with a number of 32 poles, the detrended series T_1 and T_2, have been obtain. Figure 4 shows the changes of the estimates of the correlation dimensionalities with *m* for the detrended series T_1 and T_2. The delay time was $\tau_3 = 1$(day). These results suggest that dimensionality of the time series T_1 has a good convergence with *m.* It reaches a plateau of ~4.1 for $m \geq 7$. Compared to series T_1, the dimensionality of series T_2 does not seem to reach a plateau, even though there is an obvious tendency to converge with *m* to the value ≈ 5.1. It seems that the correlation dimension of the entire Arosa time series (T_1) is decreased by the presence of the two consecutive years mentioned above having as daily values the corresponding multiannual averages. The same conclusion can be drawn if the differentiation (D) is used to detrend the time series. Since the differentiation tends to accentuate the noise, it is necessary to remove it before applying the differentiation procedure. In order to remove the noise, the singular value decomposition method with 9 × 9 correlation matrix was used

(Vautard, Ghil, 1989). Figure 4 shows the estimates of the correlation dimensionalities as function of m for the series T_1 and T_2 with noise removed and detrended by the differentiation procedure and a delay time $\tau_1 = 1$(day). In this case, the dimensionality of series T_1 reaches a plateau of ~ 4.2 for $m \geq 7$, and series T_2 has a clear tendency to converge with m to the value $d_c \approx 5.1$. The fact that the use of two different procedures to detrend the data series T_1 and T_2 results in the same values for the attractor dimensionalities, could be an argument in favour of the idea that the correlation dimension of the entire Arosa time series (T_1) is decreased by the presence of the two consecutive years having as daily values the corresponding multiannual averages, and $d_c > 5$ is a more realistic value of the attractor dimensionality, in accordance with the results obtained by Yang *et al.* (1994).

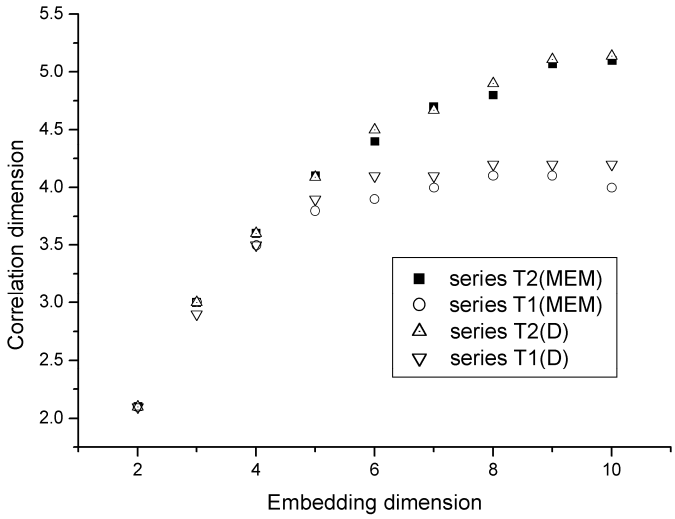

Fig. 4 – Correlation dimension versus embedding dimension for the series T_1 and T_2.

The entropy of time series T_1 and T_2 detrended by making use of MEM changes from 0.687 to 0.463, and from 0.528 to 564, respectively, when $m \in (5,9)$.

When series T_1 and T_2 are detrended by differentiation, the entropy varies from 0.485 to 0.455, and from 0.535 to 0.562, respectively, when $m \in (5,9)$.These entropy values result in a time interval of 1–2 days on which a deterministic prediction can be made.

Lyapunov exponents have been computed for embedding dimension $m = 6$ (which should be somewhat higher than the expected dimension of the attractor), a number of three sample intervals over which each pair of points is followed, and a relative accuracy of 10^{-4}. The exponent value was 0.245 ± 0.006 (d^{-1}) for both series T_1 and T_2.

The values of the correlation dimension, entropy and Lyapunov exponents discussed above demonstrate that the total ozone system is characterized by a strange attractor. Its dimensionality indicates that the essential variables which determine the total ozone dynamics in the atmosphere is between 6 and 13.

4.4. SURROGATE DATA TESTING

In order to certify the existence of the total ozone attractor of low dimensionality, the surrogate data test has to be applied. The surrogate data are obtained by computing the Fourier transform of the original data, then by generating a set of random phases in the interval $[0,2\pi]$ and finally by computing the inverse transform using the original amplitudes and the set of random phases. The surrogate series has the same spectral properties as the original one but with the determinism removed. Any statistics computed for the original data (*e.g.*, correlation dimension) should be computed for the surrogate series as well. Analysis of the surrogate data should provide values that are significantly different of those derived from the original data. If this is the case, then the null hypothesis the surrogates are consistent with can be rejected. In order to be sure that the difference is statistically significant, many surrogate data sets have to be generated and to see whether the results from the original time series lie within the range of values corresponding to the surrogates. If they do, then the difference is not statistically significant and the original data is indistinguishable from colored (correlated) noise. Figures 5 and 6 show the correlation dimension versus m for the surrogate data as compared with the detrended series T_1 and T_2. These figures suggest that, unlike the correlation dimensions for the detrended series, the correlation dimensions of the surrogate series has a quite different behavior and no longer exhibits a scaling region. Using 19 surrogate data sets, the null hypothesis can be rejected at a 95% significance level. This proves that determinism is present in the Arosa time series and the reliability of the dimension estimates of the underlying dynamical system is confirmed.

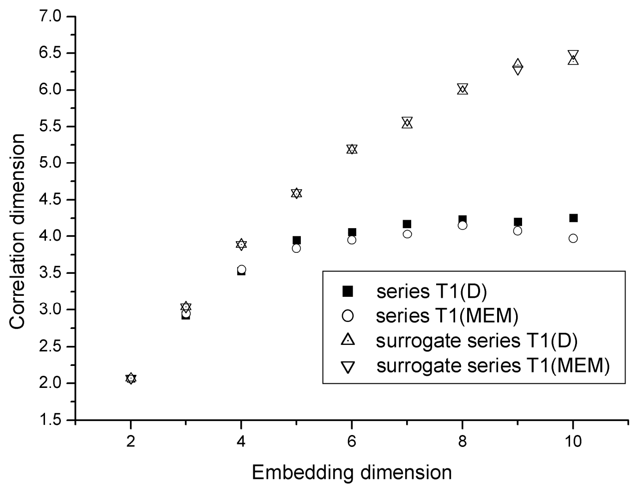

Fig. 5 – Surrogate data test for the series T_1.

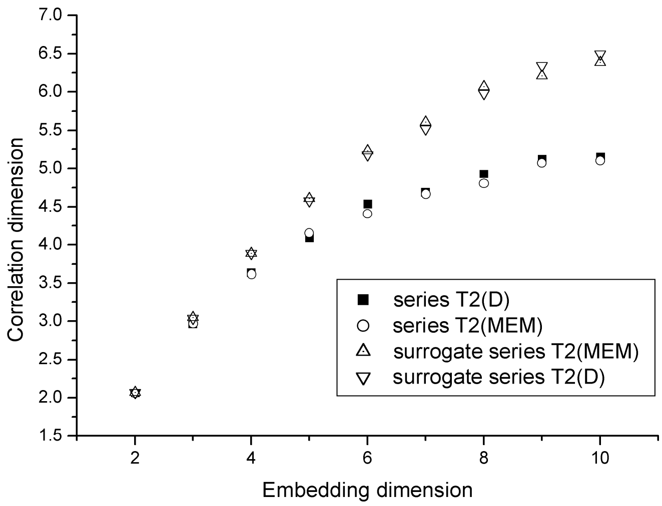

Fig. 6 – Surrogate data test for the series T_2.

5. DISCUSSION AND CONCLUSIONS

In this paper a dynamical systems analysis has been performed for the Arosa time series with the missing data filled in with the multiannual averages obtained from the daily values measured within the period 1926−2005. Taking into consideration the fact that, for a period of about two consecutive years, the daily values have been filled in with the same multiannual averages, in order to have the possibility to quantify the influence of this procedure on the characteristics describing the dynamics of the entire series, a time series for the period 1931−2005 has been considered, too, by excluding the time interval 1926−1930. The broad band spectrum exhibits a dominant frequency corresponding to seasonal variability of daily averages of the total ozone amount. Because of nonstationarity, both time series have been detrended before estimating the attractor dimensionality, by making use of the maximum entropy method and differentiation procedure. The results of both techniques corroborated with the surrogate data test indicated almost identical values (4.1 and 4.2) for the saturated correlation dimension of the Arosa time series covering the period 1926−2005. The times series for the period 1931−2005 does not reach a plateau but has a clear tendency to converge with m to a value of ≈ 5.1. The existence of the two years with the multiannual averages as daily values actually reduces the dimensionality of the ozone system attractor by about 1.

The correlation dimension of the Arosa series attractor seems to be greater than 5.

In order to be able to obtain a reliable estimate of d_c, the number of data points has to be greater than the present number of data. For example, in case of a time series having 32,000 data points, the largest correlation dimension that can be reliably determined is about 5, for which an embedding dimension of 10 should suffice. Any dimension that one would calculate for such a time series in higher embedding dimension would almost surely be spurious (Sprott, 2003; Tsonis, 1992). The attractor dimensionality of the Arosa time series indicates that the number of essential variables which determine the total ozone dynamics in the atmosphere appears to be between 6 and 13. One explanation for the low number of degrees of freedom necessary to describe the variability of the ozone system is the fact that the Dobson spectrometer provides only the vertically integrated amount of ozone, no information being available on the vertical structure of the ozone concentration.

Also, a value of $d_c > 5$ for the total ozone time series is consistent with the dimensionalities of the strange attractors of the atmospheric parameters, such as, wind, temperature and pressure (Zeng *et al.*, 1992). In addition, the fully developed turbulence is treated as a physical process of such dimensionality. According to Lorenz (1991) the dimensionality of a dynamical system attractor is determined by strong coupling of variables in the system. As regards the atmosphere, it might be viewed as a loosely coupled set of lower-dimensional subsystems and the determined dimension could be associated only with one of the subsystems. The dimensionalities computed in this paper should thus be considered as the lower limit of the number of variables which most strongly influence the ozone subsystem variability (Yang *et al.*, 1994). The fractal correlation dimensions and the positive values of the Lyapunov exponent and entropy reveal the chaotic dynamics of the Arosa time series. According to the entropy values, a deterministic prediction of the total ozone column can be done for a time interval of 1 to 2 days. Beyond this interval only a statistical prediction makes sense.

REFERENCES

GHIL, M. *et al.* (2002), *Climatic time series analysis*, Reviews of Geophysics, **40**, 1.

GRASSBERGER, P., PROCACCIA, I. (1983), *Measuring the strangeness of strange attractors*, Physica, D 9, 189−208.

JANOSI, I.M., MULLER, R. (2005), *Empirical mode decomposition and correlation properties of long daily ozone records*, Physical Review, E 71, 056126.

KANTZ, H., SCHREIBER, T. (1997), *Nonlinear Time Series Analysis*, Cambridge Univ. Press, New York.

LORENZ, E.N. (1991), *Dimensionality of weather and climate attractors*, Nature, **353**, 241−244.

SPROTT, J.C. (2003), *Chaos and Time-Series Analysis*, Oxford University Press, Oxford.
SPROTT, J.C., ROWLANDS, G. (1995), *Chaos data Analyzer − The Professional Version, Physics Academic Software*, American Institute of Physics, New York.
TAKENS, F. (1981), *Detecting strange attractors in turbulence*, Vol. 898 of *Lecture Notes in Mathematics*, edited by D.A. Rand and L.-S. Young, 366−381, Springer-Verlag, New York.
TSONIS, A. (1992), *Chaos: From Theory to Applications*, Plenum, New York.
VAUTARD, R., GHIL, M. (1989), *Singular spectrum analysis − A toolkit for short noisy chaotic signals*, Physica, D, **35**, 395−424.
WOLF, A., SWIFT, J.B., SWINNEY, H.L., VASTAND, J.A. (1985), *Determining Lyapunov exponents from a time series*, Physica, 16D, 285–317.
YANG, P., BRASSEUR, G.P., GILLE, J.C., MADRONICH, S. (1994), *Dimensionalities of ozone attractors and their global distribution*, Physica, D 76 , 331−343.
ZENG, X., PIELKE, R.A. (1993), *What does a low-dimensional weather attractor mean?* Phys. Lett. A., **175**, 299−304.
ZENG, X., PIELKE, R.A., EYKHOLT, R. (1992) *Estimating the fractal dimensions and predictability of the atmosphere*, J. Atmos. Sci., **49**, 649−659.
http //www.iac.ethz.ch/en/research/chemie/tpeter/totozon.html

Received: November 8, 2007
Accepted for publication: February 14, 2008