Comments on "Dimension Analysis of Climatic Data"

A. A. TSONIS AND J. B. ELSNER

Department of Geosciences, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin

7 November 1989

ABSTRACT

In a recent paper Mohan et al. presented a reanalysis of climatic data using concepts from the theory of dynamical systems. The data is the oxygen isotope ratio $^{18}O/^{16}O$ record of the V28–238 deep sea core covering a period of a million years at a sampling time of 2 Kiloyears. This dataset was first analysed by Nicolis and Nicolis who reported that the dynamics of the records may be explained by a low-dimensional dynamical system. We take this opportunity to bring to the attention of the scientific community some major problems involved with the reanalysis of the data hoping that this comment will serve as a reference for other analyses of different datasets in the future.

1. Introduction

According to the theory of dynamical systems, the best way to study the dynamics of a system is via the *state space.* The state space is a coordinate system whose coordinates are the variables that describe the system. At each time step the state of the system can be represented by a point in the state space. By connecting these points a trajectory that describes the evolution of the system is defined. This trajectory converges on the *attractors,* which describe the asymptotic final state of the system. The attractors may be simple topological structures such as a point, a limit cycle, or a torus; or they can be nontopological submanifolds characterized by fractal geometry (see for example Tsonis and Elsner 1989).

If the mathematical description of a dynamical system is given, the number of variables is known and the generation of the state space and of the attractor is straightforward.

If the mathematical formulation of a system is not available, the state space can be replaced by the *phase space.* The phase space may be produced using a single record of some observable variable $x(t)$ from that system (Packard et al. 1980; Reulle 1981; Takens 1981), using $x(t)$ and its successive shifts as the coordinates. Thus, given an observable $x(t)$, one can generate the complete state vector $\mathbf{X}(t)$ by using $x(t+\tau)$ as the first coordinate, $x(t+2\tau)$ as the second coordinate, and $x(t+n\tau)$ as the last coordinate. Here τ is a suitable delay parameter. This way we can define the coordinates of the phase space, which should approximate the dynamics of the system from which the observable $x(t)$ was sampled (or in other words the unknown state space). The parameter n is often referred to as the embedding dimension. For an n-dimensional phase space, a "cloud" or a set of points will be generated. The Hausdorff–Besicovitch dimension of this set can be estimated by covering the set by n-dimensional cubes of side length l and determining the number of cubes $N(l)$ needed to cover the set in the limit as l goes to zero (Mandelbrot 1983). This is the *box-counting algorithm* and if this number scales as

$$N(l)\,\alpha\, l^{-d} \qquad l \to 0 \tag{1}$$

then the scaling exponent d is an estimation of the Hausdorff–Besicovitch dimension for that n. In a $\log N(l)$ vs. $\log l$ plot, the exponent d can be estimated by the slope of a straight line (the scaling region). Using the state vector $\mathbf{X}(t)$ we can test Eq. (1) for increasing values of n. If the original time series is random, then $d = n$ for any n (a random process embedded in a n-dimensional space always fills that space). If, however, the value of d becomes independent of n (that is, reaches a saturation value D_0), it means that the system represented by the time series has some structure and should possess an attractor whose Hausdorff–Besicovitch dimension is equal to D_0. The above procedure for estimating D_0 is a consequence of the fact that the actual number of variables present in the evolution of the system is not known and thus we do not know a priori what n should be. We must, therefore, vary n until we "tune" to a structure that becomes invariant in higher embedding dimensions (an indication that extra variables are not needed to explain the dynamics of the system in question).

The above numerical approach to estimate the di-

Corresponding author address: Dr. Anastasios A. Tsonis, Dept. of Geosciences, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201.

mension of an attractor from a time series is, however, very limited. The reason for that is that an enormous number of points on the attractor is required to make sure that a given area in the phase space is indeed empty and not simply visited rarely. It has been documented (Froehling et al. 1981; Greenside et al. 1982) that a box-counting approach is not feasible for phase space dimensions greater than two.

An alternative approach, which is much more applicable, has been developed by Grassberger and Procaccia (1983a, 1983b). This approach again generates in an n-dimensional phase space a cloud of points. But instead of covering the set with hypercubes, it finds the number of pairs $N(r, n)$ with distances less than a distance r. In this case, if for significantly small r, we find that

$$N(r, n) \alpha r^{d_2}. \tag{2}$$

The scaling exponent d_2 is the correlation dimension of the attractor for that n. We then test Eq. (2) for increasing values of n and check, as previously done, for a saturation value D_2, which will be an estimation of the correlation dimension of the attractor. It should be mentioned at this point that τ can be small, but care should be taken not to include in the sums pairs whose time separation is less than the correlation time. The correlation dimension D_2 is less than the Hausdorff–Besicovitch (or fractal) dimension D_0 and actually measures the spatial correlation of the points that lie on the attractor. For a random time series there will be no such spatial correlation in any embedding dimension and thus no saturation will be observed in the exponent d_2. The above approach still requires a large number of points (especially for high embedding dimensions), but at least it is more feasible than the box-counting method. Thus, most of the analyses up to date [including the analysis of Mohan et al. (1990; MRR)] have concentrated in calculating D_2. It should be mentioned here that one very important consequence of knowing the Hausdorff–Besicovitch or any other dimension of an attractor is that the dimensionality of an attractor, whether fractal or not, indicates the minimum number of variables present in the evolution of the corresponding dynamical system (in other words the attractor must be embedded in a state space of at least its dimension). Therefore, the determination of the Hausdorff–Besicovitch dimension (or for that matter of any other generalized dimension) of an attractor sets a number of constraints that should be satisfied by a model used to predict the evolution of the system.

2. Critical embedding dimension

The problems begin when researchers use experimental data to reconstruct the attractor, causing the existence of the scaling region to depend greatly on the amount of data. As we will try to argue next, given a certain amount of data there exists a critical embedding dimension, n_c, beyond which the procedure for estimating any dimension is invalidated and should not be attempted.

Because of Eq. (2), the population of pairs of points on smaller scales is smaller than the population of pairs on longer scales. Thus, if for a fixed n the number of points in the set becomes smaller, the population of pairs over the scales for which Eq. (2) holds begins to be depleted. As we continue to decrease the number of points (for our fixed n always) we will observe:

a) more and more depletion at smaller scales (since less and less points will be found) and

b) large fluctuations of $N(r, n)$ due to small populations at larger scales.

The net result is that the scaling region may be completely masked. Any straight line fitting at this point will result in a false correlation dimension for that n. Thus for an accurate estimation of the slope d_2 on a $\log N(r, n)$ versus $\log r$ plot requires a minimum number of points.

It should now be emphasized that by embedding the dataset into continually higher dimensions we effectively "distribute" the same number of points into continually higher dimensional space. In effect we go from a densely populated low-dimensional space to a sparsely occupied high-dimensional space. The result will be the same as before: for very small scales $N(r, n)$ goes to zero (depopulation) and for large scales (scales close to the radius of the set) $N(r, n) = N$ independently of r and n(saturation). Thus, at some embedding dimension the scaling region will not be clearly defined as it will be "lost" between depopulation and saturation. The embedding dimension above which the scaling region cannot be accurately defined is called the critical embedding dimension n_c. The above points have been emphasized by Essex et al. (1987), but unfortunately have not been seriously considered in several reported climate and weather attractor reconstructions including that of MRR.

We now wish to demonstrate the above by starting with a time series of 500 white noise values. We know that in this case for any embedding dimension $d_2 = n$ (as long as we use the necessary number of points for each n). We started with embedding dimension $n = 2$ and found $\log N(r, n)$ as a function of $\log r$ and then we calculated

$$\text{slope} = \Delta \log N(r, n)/\Delta \log r$$

as a function of $\log r$. If there exists a clearly defined scaling region in the $\log N(r, n)$ versus $\log r$ plots then we should be able, on a slope versus $\log r$ plot, to observe a plateau. This plateau will provide an estimation for the exponent d_2 for a given n. Figure 1 shows slope versus $\log r$ for $n = 2$ (triangles), $n = 6$ (squares), and $n = 15$ (stars). For $n = 2$ we observe that slope is nearly constant at about a value of two for a wide range of

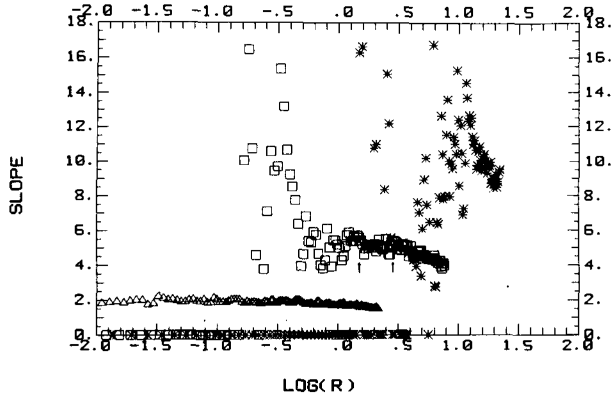

FIG. 1. $\Delta \log N(r, n)/\Delta \log r$ as a function of $\log r$ for embedding dimension 2 (△), 6 (□), and 15 (∗). Based on a record of 500 white noise values. See text for details.

scales. When the scales become too large, saturation indicated by a gradual decrease of slope is seen. Depopulation is not visible (at least within the scale range of the figure). Therefore we may conclude that 500 points are adequate in defining the scaling region when $n = 2$. For $n = 6$ we observe a very different picture. We see depopulation manifesting itself as many zero slope values over small scales, large fluctuations over larger scales, and saturation over very large scales. A scaling region can be suggested (indicated by the arrows) but it is not as clearly defined or as wide as in the case of $n = 2$. Nevertheless this small plateau is found at about slope = 5.0, which is less than the true value of 6.0. Thus an attempt here to define a scaling region will at best result in an underestimation of the true value of the exponent d_2. Similar comments can be made for $n = 15$; here the difference shows that there is virtually no way one can define a scaling region.

Figure 2a is similar to Fig. 1 but for $n = 4$. Similar comments to those made in Fig. 1 for n greater or equal to 6 can be made for Fig. 2a. A scaling region (indicated by the arrows) may be identified but it will produce a value of $d_2 = 3.5$ (which is less than the true value of 4.0). However, in Fig. 2b (with 5000 points), a well-defined scaling region exists with a value of $d_2 = 4.0$. It is important to note that the true scaling region here is *at smaller* scales compared to the scaling region that is identified when 500 points are used (where the true scaling region is masked by large fluctuations).

The above figures clearly demonstrate what was discussed previously. We should be very careful not to exceed the critical embedding dimension that is a function of the data size. In fact with 500 data points the scaling region is ill-defined and any estimated value of the exponent d_2 is false for n greater that 3.0. Therefore it should not be attempted nor reported.

We find no justification at all in MRR for going from 512 points up to embedding dimension $n = 15$ and actually estimating a slope. As can be seen from Figs. 1 and 2, the scaling region is not found in a fixed interval of scales. In general as n increases the $\log N(r, n)$ versus $\log r$ plot is shifted towards higher r values. It is shown that even if the scaling region exists for any n, it does not exist over the same scale range for all embedding dimensions. Thus it is not appropriate to estimate d_2 by considering the points that fall within a fixed interval in $\log r$ throughout a $\log N(r, n)$ versus $\log r$ plot. The scaling region, and consequently the slope, should be estimated using figures like our Fig. 1 or Fig. 2. Thus we believe that all the slope values reported by MRR in Figs. 2 and 3, as well as the accompanied interpretations and conclusions, are in error. Moreover, it is particularly confusing that MRR state in section 3 and in their conclusions that it is not possible to embed the data in more than three dimensions and obtain meaningful results; and yet they go ahead and present results and conclusions by embedding the data up to an embedding dimension $n = 15$. In the initial analysis Nicolis and Nicolis (1984) went up to embedding dimension six and used a random sample to test the significance of their results.

3. Conclusions

Although MRR raise some interesting points their reanalysis may confuse the scientific community about

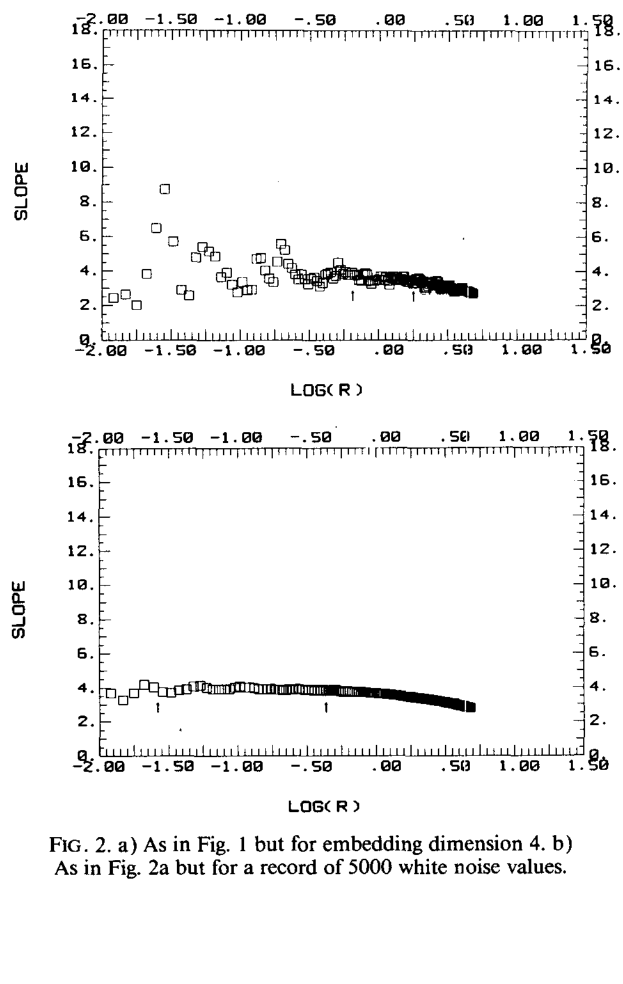

FIG. 2. a) As in Fig. 1 but for embedding dimension 4. b) As in Fig. 2a but for a record of 5000 white noise values.

the usefulness and the correct methods used in the search for attractors in climatic and/or weather observables. The theory of dynamical systems and chaos has provided us with new tools in analyzing observables. Whether or not low-dimensional attractors exist in nature is still debatable. Thus it is very critical that we apply these new theories with the utmost of care. Otherwise we may damage a theory that right now shows a great deal of promise.

REFERENCES

Essex, C., T. Lookman and M. A. H. Nerenberg, 1987: The climate attractor over short timescales. *Nature,* **326,** 64–66.

Froehling, H., J. P. Crutchfield, D. Farmer, N. H. Packard and R. Shaw, 1981: On determining the dimension of chaotic flows. *Physica,* **3D,** 605–617.

Grassberger, P., and I. Procaccia, 1983a: Characterization of strange attractors. *Phys. Rev. Lett.,* **50,** 346–349.

——, 1983b: Measuring the strangeness of strange attractors. *Physica,* **9D,** 189–208.

Greenside, H. S., A. Wolf, J. Swift and T. Pignarato, 1982: The impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors. *Phys. Rev.,* **A25,** 3453–3456.

Mohan, T. R. K., J. S. Rao and R. Ramaswamy, 1989: Dimension analysis of climatic data. *J. Climate,* **2,** 1047–1057.

Nicolis, C., and G. Nicolis, 1984: Is there a climatic attractor? *Nature,* **311,** 529–532.

Packard, N. H., J. P. Crutchfield, J. D. Farmer and R. S. Shaw, 1980: Geometry from a time series. *Phys. Rev. Lett.,* **45,** 712–716.

Ruelle, D., 1981: Chemical kinetics and differential dynamical systems. *Nonlinear phenomena in chemical dynamics,* A. Pacault and C. Vidal, Eds., Springer-Verlag, 30–37.

Takens, F., 1981: *Dynamical systems and turbulence. Lecture notes in Mathematics.* Springer, 366–381.

Tsonis, A. A., and J. B. Elsner, 1989: Chaos, strange attractors, and weather. *Bull. Amer. Meteor. Soc.,* **70,** 14–23.