Tellus (1995), 47A, 457–472
Printed in Belgium – all rights reserved

TELLUS
ISSN 0280-6495

Degrees of freedom in Northern Hemisphere circulation data

By ZOLTAN TOTH*, *NOAA/NMC, Development Division, 5200 Auth Rd., Room 204, Camp Springs, MD 20746, USA*

(Manuscript received 2 May 1994; in final form 18 October 1994)

ABSTRACT

Estimates of the number of degrees of freedom (dof) of the 700 hPa Northern Hemisphere extratropical wintertime daily and time average circulation patterns are presented using two statistical methods. In the first method, it is assumed that the distribution of the mean square distances between the circulation patterns and the climate mean follows a χ^2 distribution. The number of dof of the best χ^2 fit to the empirical data is the most probable number of dof of the circulation data. The other, newly developed hyperspheres method compares the local density properties in the empirical data to density in independently and identically distributed (iid) normal distributions with different number of dof. The best fit, again, determines the most probable number of dof for the observed data. Though the hyperspheres method uses the smallest scales available in the data, just as the Grassberger-Procaccia method (GPM) and other "local" dimension estimation methods, it has distinct features. Most importantly, the GPM assumes that the empirical data is unbounded and distributed uniformly in the multivariate space. In contrast, the hyperspheres method, in accordance with many geophysical applications, assumes a bounded distribution in which density is a strong function of distance from the mean. Additionally, the hyperspheres method directly uses density information, that is neglected by GPM. The advantages of the hyperspheres method as compared to GPM are: (1) no systematic negative bias is present in the dof estimates; (2) the edge effect is eliminated; (3) no scaling is necessary. The results from the two methods overlap, thus indicating that at the 10% significance level the number of dof of the daily circulation data set is 24. This result pertains to a finite dataset in which the definitional requirement for dynamical dimension that the scales measured go to zero is not satisfied. We interpret our dof results as the dimension of a hypothetical subset of the atmospheric circulation that governs the large scale motions. Since the database of this study covers only part of the global atmosphere, the dimension of the whole atmosphere would be larger than our estimate. Our qualitative dimension estimate for the global circulation (60–90) refers to the minimum number of independent variables needed to model the large scale, low frequency variability of the atmosphere. In the present study grid point values served to describe the dynamical system of the atmosphere (spatial embedding, instead of time-delay or temporal embedding.) An argument is made that due to weak coupling between remote processes, estimates with temporal embedding measure only the dimension of a regional subset of the large-scale global circulation.

1. Introduction

There have been an increasing number of studies concerned with the dimensionality of atmospheric motions. The motivation behind these studies is at least twofold. First, it is a theoretically interesting question. Although, as Vautard and Ghil (1989) noted, the assessment of the number of degrees of freedom (dof) or independent dimensions is only a first step in understanding the real complexity of a problem, it is a challenge to researchers. Another and more practical part of the motivation is that, to perform certain statistical analyses or tests (see, e.g., Livezey and Chen, 1983), an a priori

* Current affiliation: General Sciences Corporation, Laurel, MD, USA.

knowledge of the number of independent variables (or dof) in the data set in question is needed.

When one estimates the dimensionality of a dynamical system such as the atmosphere, the system can be represented by two types of data sets (or by a combination of them.) First, the phase space of a system can be described by different variables measured at a sufficiently large number of spatial locations of the system (spatial embedding). Secondly, the phase space can be reconstructed by time-delay or "temporal" embedding. The theory behind temporal embedding (see, e.g., Takens, 1981) guarantees that the phase space of a *dynamical system* (a system that can be represented by a finite number of coupled ordinary differential equations) can be reconstructed, with very few exceptions, from a single-state variable. In this case, the full phase space of the system is embedded into an n-dimensional space spanned by the time series of an arbitrarily selected variable and its $n-1$ time lag series. Here n should be sufficiently larger than D, the dimension of the system ($n > 2D + 1$). The obvious advantage of temporal embedding is that it does not require the knowledge of many different variables to describe the phase space of a system.

Notice that with either spatial or temporal embedding, the system is described in an embedding space whose number of variables n is higher than the independent variables in the system itself. In either case, once we have a full representation of the system, we can attempt to estimate the dof present in the actual data set. Throughout the paper we will refer to these estimates, based on limited observational data, as "dof", in contrast to "dimension" that is a property of the underlying object, the dynamical system itself. As we will see later it is not obvious if dof can be interpreted as the dimesion of the underlying object or only as a statistical property of the data set at hand.

There exist two groups of methods to estimate the dof of data sets. In the estimation, the global methods use information from all scales present in the data set, while the "local" methods focus on the smallest phase space scales present in the data. The next two subsections offer a short review of the basic methods and earlier results.

1.1. Global methods

Long before the theory about dynamical systems and temporal embedding developed, the dof in atmospheric circulation data had been estimated based on spatial embedding. Lorenz (1969) studied daily NH extratropical circulation data combined from three different levels of the atmosphere. He computed a mean square type distance (msd, see eq. (5) below) between all possible pairs of circulation patterns in his data base. He assumed that the difference between two patterns at each grid point follows a normal distribution. Hence the averaged sum of the square of these differences (msd) should follow a χ^2 distribution. Comparing the empirical distribution of msd's to χ^2 distributions with different dof, he got the best fit at 44 dof, thus concluding that the above data set has approximately that many dof. The χ^2 method described in Subsection 3.1 is a close variant of Lorenz' method. Fraedrich et al. (1994), with a generalized version of this method, arrived at dof estimates between 13 and 36 for midlatitude NH unfiltered and low-pass filtered 1000 hPa height data.

Livezey and Chen (1983), as a by-product of their pioneering field significance study, presented an estimate of 35 for the dof of the 700 hPa Northern Hemisphere wintertime monthly mean circulation field. They also noted that the question of spatial dof can be viewed similarly to that of an integral time scale. Barnett and Preisendorfer (1987) followed this idea when computing the effective number of independent stations in their United States surface temperature data set.

An approach somewhat similar to that of Lorenz (1969) but based on pattern correlations was proposed by Van den Dool (personal communication, 1990). It uses a well known relationship between the standard deviation (sd) of a linear correlation and dof (see, e.g., Panofsky and Brier, 1968):

$$\mathrm{sd} = (\mathrm{dof} - 2)^{-1/2}. \tag{1}$$

Eq. (1) is valid if the correlation is normally distributed with an expected value of 0. Horel (1985) applied this method for pattern correlations between daily 500 hPa extratropical NH height data and found that the unfiltered data have 30–37 dof. When all wave numbers greater than 4 were eliminated, the dof dropped to 21–27. Van den Dool and Chervin (1986) studied outputs of the Community Climate Model. They found that the number of dof of monthly averaged 500 hPa

(700 hPa) forecast height fields over the extratropical NH is around 20–21 (18), while for the whole globe it is 32–33. For observed wintertime monthly mean 500 hPa height fields, their estimate ranges between 14 and 27.

None of the above methods assessing the dof gives indications of the real independent directions in a data set at hand. By a principal component analysis (PCA) we can determine the directions which maximize the variance of the data. Moreover, as O'Lenic and Livezey (1988) showed, a careful study of the principal components, with different number of modes retained for rotation, offers a tool for separating the noise from the rest of the variance in the data. The number of eigenvectors whose eigenvalues are statistically distinguishable from the flat tail of the spectrum (that is presumably associated with noise) can be considered as another estimate of the dof. Studying 700 hPa low-pass filtered data, they concluded that much of the meteorologically meaningful variance can be reproduced by retaining only 13 eigenmodes for rotation. However, when interpreting and comparing their results to those from other methods we have to note two differences. First, the principal components explain a decreasing amount of variance in the data with increasing mode number while the other methods assume that each independent variable has the same variance. Consequently, in this respect pca (and, to a lesser extent, rotated pca) gives an underestimation of dof defined in terms of independent variables with equal variance. Secondly, pca (as well as the correlation based method above) is based on linear correlations. If there are nonlinear connections between the variables, the dimension estimate will not reflect these connections and hence will be inflated.

1.2. Local methods

As we will see below, dynamical dimension is defined in a limit sense, where the scales considered in an object go to zero. Global methods that do not differentiate among different scales are not the best tools in this case, unless the structure of the attractor does not change with different scales, i.e., it is *self-similar*. In a general case, one needs a method that enables us to evaluate dof with respect to the smallest phase space scales available in the observed data. The definitions and description below is given after Fraedrich (1986), Grassberger and Procaccia (1983) and Henderson and Wells (1988), wherein further references can be found.

If we fill a geometrical object of a volume V by a number $N(L)$ of smaller "boxes" with sidelength L, the volume of the object can be written as

$$V = \sum_{N(L)} L^{D_0} = N(L)\, L^{D_0}, \tag{2}$$

where D_0 is the dimension of the phase space containing the object. The number of boxes needed to fill the volume increases in proportion to the inverse power D_0 of the sidelength. Theoretically, by decreasing L one can assess the "box" or "capacity" dimension of any system. As seen from eq. (2), D_0 can be noninteger. If the object is a description of a dynamical system, this would be an indication that the system, though deterministic, has a chaotic behavior (as compared to a system with integer D_0, characterized by (quasi) periodic behavior). The next larger integer than D_0 in (2) can be interpreted as the *minimum* number of variables needed to describe the evolution of the system in phase space.

The above box counting method, however, converges very slowly. Therefore another method has been suggested by Grassberger and Procaccia (1983) to estimate the lower bound of dimensionality, that is called "correlation dimension" (D_c or D_2). This method is based on counting distances between randomly chosen pairs of points on the attractor. Let us consider the number of pairs $N(l)$ within a critical distance l. In this case, if the number of points (N) goes to infinity, then the cumulative distribution function $C(l)$, after it is normalized by the total of N^2 pairs, is:

$$C(l) = \frac{\lim}{N \to \infty} N(l)/N^2 \sim l^{D_c}. \tag{3}$$

As an example, consider data points homogeneously distributed on a line (2-dimensional surface/3-dimensional volume). The number of pairs closer than l grows linearly (quadratically/cubically) with increasing l, or, in other words, proportionally to l (l^2/l^3). In practice, both N and l are limited by the finite observational samples. Thus, one has to record $C(l)$ as a function of l, and D_c is determined as the slope of l versus $C(l)$ on a logarithmic plot (*scaling*). This is an extrapolation, based on the assumption that the underlying fractal is statistically self-similar, i.e., the attractor's structure is invariant with respect to decreasing scales. It is shown in Grassberger and Procaccia (1983) that the correlation dimension D_c derived from (3) is closely related to, but never greater than, the fractal dimension defined by (2).

The Grassberger-Procaccia method (GPM) has been extensively applied to different data sets to assess the dimension of the "climate" or "weather" attractor. Almost exclusively, temporal embedding is used in these studies. However, it is not clear what the advantage is, if any, of using temporal embedding when applied on atmospheric circulation data (in contrast to paleoclimatic data where spatial coverage is scarce.) In this case a reliable set of different variables with a good spatial resolution can be used, as we saw earlier, for spatial (or a combination of spatial and temporal) embedding.

Among the studies using temporal embedding, we first refer to Fraedrich (1986, 1987), who investigated daily sea level pressure, relative sunshine duration and NH zonal wavenumber five amplitude series and after removing the interannual variability and seasonal changes he found the correlation dimension between three and seven. Essex et al. (1987) studied 500 hPa height series at nine nearby locations over Northern America. Combining the nine data series they concluded that the climate attractor has a correlation dimension less than 7. However, the authors were not confident about the interpretation of their results; whether the correlation dimension is a geographically local or global characteristic. Henderson and Wells (1988), perhaps to avoid such a question, investigated time series of the difference between the zonally averaged 500 hPa height at 60° N and 30° N. However, their fractal dimension estimate (6–7) is also for a geographically confined (and spatially averaged) variable. Keppenne and Nicolis (1989) studied 500 hPa height data series from 9 stations over Western Europe. Their correlation dimension estimates for individual stations are about 7–8. From this fact they concluded that the 9 time series "refer to a well-defined dynamical system describing the short-term variability of the Western European weather." However, it is not clear what are the exact geographical or other specifications of such a hypothesized dynamical system. Note that at other parts of the hemisphere (Essex et al., 1987) similar estimates were derived.

Some of the above and other related studies received substantial criticism in the literature (see, e.g., Ruelle, 1990), partly because of the limited amount of data used. It suffices to say here that the interpretation of dof and dimension estimates is a question open to debate.

1.3. Open questions

One of the major issues addressed in the current study is the apparent difference between dof estimates from global versus local dimension estimation techniques. The former estimates are in the range of 13–44 while the latter in the range of 3–8. Two logical explanations offer themselves. First, it is possible that the global methods, relying mainly on medium and larger scales in the data set, do not detect important connections present among different variables in the data. If these connections are present only at the smaller scales in the phase space, the dof estimate from global methods may result in much larger numbers than those from local methods, which focus specifically on the small scales. According to the second possible explanation, the difference between the two types of estimates may be a result of the difference in embedding. All global method applications used spatial embedding while the local applications used temporal embedding. However, as Lorenz (1991) pointed out, in case of weakly coupled variables, temporal embedding may result in a spuriously low dimension estimate. To identify which explanation has more relevance, we will apply the same, global method (a variant of that used by Lorenz, 1969, see Subsection 3.1) with both temporal and spatial embedding.

Another major issue discussed in this paper is some inherent limitations present in the GP and related local methods. As Grassberger and Procaccia (1983) showed, the correlation dimension defined by them gives an underestimation of the real fractal dimension. We can easily show this if we return to the example of randomly scattered points on a line. In this case the number of points in a certain interval is proportional to the length of the interval only if the points are distributed *uniformly*. If the data points are distributed, for example, normally, the GP method will yield a dimension smaller than one. This is because except at the tails of the distribution (i.e., outside the two inflection points of the density curve), if we increase the length of the interval (l), the number of points within the interval increases, on average, slower than that expected from a linear relationship. Since the majority of the points lies inside the inflection points, the effect is a decrease in the dimension estimate. In more than one dimension, the effect of non-uniformity is similar, leading to an underestimation of the true fractal dimension.

Other undesired consequences of finite and non-uniformly distributed samples on GPM correlation dimension estimates were noted, for example, by Essex et al. (1987) and Henderson and Wells (1988). When the counting of datapoint pairs within a certain distance l is performed, the correlation dimension can be interpreted point by point. The different values for individual points are later averaged to arrive at an integral correlation dimension estimate. The first problem is that, no matter what the statistical distribution of the data is, the points on the "edge" of the distribution may not yield dimension estimates representative of the whole sample simply due to their marginal location. (The higher the dimension, the greater this effect will be.) A second point is that the GP method requires the iterative application of different l distance values ("scaling") to count the number of pairs of data points within that range. In a non-uniform distribution the range of l yielding assessable (non-zero and non-total) counts may be very limited. To carry out the dimension estimation it is sometimes necessary to exclude some data points from the edges of the distribution after a tedious analysis of the data (Essex et al., 1987).

The problem arising from non-uniformly distributed observational data was recognized early on. Partly to remedy this problem, Hentschel and Procaccia (1983) and Grassberger (1983) introduced a generalization of the correlation dimension (D_q). As a further extension, Halsey et al. (1986) proposed a "spectrum of scaling indices", that may give a fuller description of the multifractal nature of complex physical systems such as the atmosphere. These methods assume the existence and attempt to describe the extent of lacunarity (i.e., density "wholes") within the data. However, on the large scales in the multivariate phase space the data is still assumed uniform and unbounded. We will argue that it is this latter assumption that prevents all GPM related methods from overcoming the limitations (underestimation and problems related to the edge effect) discussed above.

To alleviate some of the problems with GPM a new local dimension estimation method will be presented in Subsection 3.2. This method is based on the assumption that the large scale distribution of the data in the phase space is bounded and is approximately iid multinormal. After describing the data (Section 2) and methods (Section 3) used in this study, we will apply both dimension estimation methods to the same, spatially embedded 700 hPa extratropical NH circulation data (Section 4). A discussion of the results and the conclusions are presented in Sections 5 and 6, respectively.

2. Data and similarity measures

2.1. Data base

The dof of the wintertime circulation data will be investigated here with both daily and time averaged data. For the daily data, we averaged the two analyses of NH 700 hPa height field compiled at the National Meteorological Center (NMC) for each day. Instead of the original 541-point grid, the 358-point nearly equal area grid (designed by Barnston and Livezey, 1987) was used. This grid covers the NH poleward of 20° N, excluding some areas over Africa and Eurasia. (When estimating gradients the data over the North Pole were also used.) To perform the statistical tests detailed in the next section we need a temporally independent data set. Hence in our data set, we conservatively retained data only on every 20th day. The first case in our set of 330 winter maps is 15 October 1949, while the last case is 5 April 1982. Each map's spatial mean height value was removed from all the original gridpoint height values and only the anomalies from the map mean were kept. If z_{ij} stands for the height at gridpoint i ($i = 1, \ldots, 358$) and at time point j ($j = 1, \ldots, 330$), then the map anomaly is

$$a_{i,j} = z_{i,j} - \frac{\sum_{i=1}^{358} z_{i,j}}{358}. \tag{4}$$

The time averaged data are also represented by map anomalies. This data set consists of time averages of daily maps for "Quasi-Stationary Periods" (QSP's) for the same winters as above, as defined in Toth (1992). QSP's have an average duration of 13 days (with a range between 5 and 40 days). The average interval between QSP's is 7 days, spreading roughly from 1 to 15 days. Since the series of 290 QSP maps from 33 winters showed a slight indication for time dependence, 17 cases were removed (see the Appendix) to remedy the problem. The remaining 273 temporally independent time average maps were then subjected to dof investigations. The motivation for the use of time averaged data is primarily practical. The dof estimates are used in a follow-up paper (Toth, 1993) that analyzes the density properties of the QSP (and also daily) circulation phase space.

It was shown earlier (Toth, 1991a, 1992) that both daily and time averaged wintertime hemispheric circulation patterns are distributed in phase space approximately according to a multinormal distribution. This was found to be true, however, only in a phase-average sense, when the density information is expressed as a function of distance from the climate mean (origin), computed by averaging the density over all phase points equidistant from the mean. In other words, the radial distribution of the circulation data is normal. We call this class of distributions "phase averaged multinormal". For a formal definition, see eq. (7) in Subsection 3.1.

Note that phase averaged multinormality does not exclude fractal-like behavior. The negative effect of low density, lacunar areas on phase average density may be offset by the positive effect of dense areas situated at the same distance from the mean as the lacunar areas but in other phase directions. A true multinormal distribution is a special case in this class where all marginal distributions are normal. As a major departure from the GPM that assumes a uniform distribution in the multivariate phase space, the multinormal macro structure of the fractal of atmospheric circulation will be reflected in the methods presented in Section 3. We further note that phase-average multinormality can be reproduced by identically and independently distributed (iid) normal distributions. Such a distribution can be uniquely defined by its mean, standard deviation and dof. The former two parameters are easily estimated from the data via traditional statistical methods while the dof estimation is the subject of the next section.

2.2. *Circulation similarity measures*

As in any other circulation study, we need to define a distance function which measures the similarity between two circulation maps. Based on earlier results (Toth, 1991b) we chose two distance functions. The first of them is the widely used Root Mean Square Difference (RMSD):

$$\mathrm{RMSD} = \left[\frac{\sum_{i=1}^{358} (a_{i,j} - a_{i,l})^2}{358} \right]^{1/2}. \tag{5}$$

The other is the Difference in Gradient of Height (DGH) as a distance function:

$$\mathrm{DGH} = \frac{\left[\begin{array}{l} \sum_{i=1}^{358} [(\Delta_x a_{i,j} - \Delta_x a_{i,l})^2 \\ \quad + (\Delta_y a_{i,j} - \Delta_y a_{i,l})^2]^{1/2} \end{array} \right]}{358}, \tag{6}$$

where $\Delta_x a_{i,j}$ is the zonal and $\Delta_y a_{i,j}$ is the meridional gradient of the height field for each grid point, estimated from grid-point height values. RMSD and DGH will be used in parallel (where applicable) in the statistical procedures. Though the fractal dimension is invariant with respect to choice of norm (e.g., Theiler, 1990), it will be instructive to see whether this is true for our dof estimates, based on finite data.

3. Statistical methods to estimate dof

The dof of the daily and time average data sets will be estimated here by two partially independent methods. Both methods are based on the assumption that the data is from a multinormal distribution with identically and independently distributed (iid) variables. As shown in the previous section, this assumption was found to be true in a phase-averaged sense. The first method considers the cumulative frequency distribution of the data in a global sense, while the second concentrates on average local density characteristics. The first method can also be a specific test of phase averaged multinormality of the circulation data (with a particular dof).

3.1. *Cumulative frequency distribution*

Let $x_1, x_2, x_3, \ldots, x_m$ be realizations of normally distributed, independent variables with unit variance and zero mean. Then the sum of the square of the variables follows a χ^2 distribution:

$$\chi^2 = x_1^2 + x_2^2 + x_3^2 + \cdots + x_m^2, \tag{7}$$

with m dof. We can standardize this χ^2 variable by dividing it by the number of variables, thus retaining only the information about the shape of the distribution:

$$\chi_s^2 = \frac{x_1^2 + x_2^2 + x_3^2 + \cdots + x_m^2}{m}. \tag{8}$$

Now, consider another set of variables, the daily (or time average) realizations of gridpoint height data from eq. (4), $a_{1,j}$, $a_{2,j}$, $a_{3,j}$, ..., $a_{n,j}$, where $n = 358$, the number of gridpoints in our data set. Let $c_i = \sum a_{i,j}/N$ be the climate mean height value at gridpoint i, where N is the total number of maps in our data set. Then $ca_{i,j} = a_{i,j} - c_i$ is the anomaly from the climate mean. If these anomalies, as assumed earlier, follow a normal distribution with approximately equal variance, then the sum of the square of the variables, after standardizing their variance (by dividing the sum by the square of the standard deviation sd^2, and, as earlier, by the number of variables, n) should approximately follow a "standardized" χ^2 distribution:

$$\chi_s^2 = \frac{ca_{1,j}^2 + ca_{2,j}^2 + ca_{3,j}^2 + \cdots + ca_{n,j}^2}{\mathrm{sd}^2\, n}, \tag{9}$$

with m dof (where $m \leqslant n$), thus providing a definition of dof. Consequently, if our assumption about the multinormality of the circulation data is correct, then the empirical distribution of the standardized sum of squares in (9) for different circulation patterns can be well approximated with a standardized χ^2 distribution.

It is worth saying that the numerator of (9) is the square distance between a circulation pattern and the climate mean; hence in this test we use eq. (5), without the square root function, to measure distances in the phase space. (Note that DGH in eq. (6) is not a root mean type function and therefore cannot be employed in this test.) We can now interpret the test as a way of checking how the circulation patterns are distributed in the phase space, in regard to their distance from the climate mean (i.e., in a phase-averaged sense). If there are only a few independent variables, then the majority of the circulation patterns will be situated close to the mean (χ^2 distributions with only a few dof are very asymmetric). But if, on the contrary, there are many independent variables in the circulation data, then in a finite sample there will be no circulation pattern close to the mean at all since with increasing dof, the χ^2 distribution assumes a more and more symmetric shape, asymptotically approaching a normal distribution, with a non-zero mean.

To check the goodness of χ^2 fit to the experimental mean square distance data derived through (9) we applied a standard χ^2 test (see, e.g., Kendall and Stuart, 1967). We opted for the use of non-equally spaced classes so that the tails of the distributions can be better resolved. For the whole time-averaged data set, 10 classes were used with hypothetical frequencies (from the χ^2 distribution) of 5%, 5%, 10%, 10%, 20%, 20%, 10%, 10%, 5%, 5%. For subsets of this data set, the number of classes was reduced to eight by merging the two pairs of 5% categories at the tails of the distribution. For the daily data set with 330 observations but a higher dimensionality we used eight classes with hypothetical frequencies of 5%, 5%, 15%, 25%, 25%, 15%, 5%, 5%.

3.2. *Local density properties*

In this test, we assume again that the data, by and large, follow a multinormal distribution and the variables have approximately equal variance (Step 1 in the flow diagram of Table 1). This is in

Table 1. *Flow chart of the dof estimate method based on local density properties* (*see text for notation*)

1. Check whether the distribution of data is phase average multinormal
2. Determine N, set k
3. Determine $r_k^{D,N}(d)$ for different D, as a function of d
4. Determine ${}_{\mathrm{sigl}}\, p_k^{D,N}$, ${}_{\mathrm{sigu}}\, p_k^{D,N}$ for different D
5. Apply $H_k^{D,N}(d)$ on real data; determine $M_k^{D,N}(p)$ for different D
6. If ${}_{\mathrm{sigl}}\, p_k^{D,N} < M_k^{D,N}(p) < {}_{\mathrm{sigu}}\, p_k^{D,N}$, D is a dof estimate at the significance level used.

contrast with GPM that implicitly assumes a uniform distribution. Specifically, we will measure the local density around each data point in our dataset that has $n \gg$ dof partially dependent variables due to spatial embedding. The measured density will then be compared to density in multinormally distributed data with different number of iid variables (marked as $D=$ dof), whose variance is equal to that in the observed data. In other words, we will search for a multinormal distribution with D independent variables that can reproduce certain statistical properties (particularly, local density) of the experimental data which have n partially dependent variables. Since tables or formulas for iid multinormal density are not readily available, the test is carried out through the use of randomly generated iid data ("surrogate" data.)

Let a denote a circulation map in the phase space that is populated with multinormally distributed data of sample size N and $r_k(a)$ denote the distance of the kth closest neighbor from the point a. In a statistical sense, $r_k(a)$ will depend on two factors. The first of these factors is the dof present in the data. It can be seen that the more dof the data have, the fewer neighboring data points there will be within a certain distance around any particular data point in the phase space. This is because, in a geometrical sense, there is more "space" for the fixed number of data points to be distributed in when there are more dimensions. The other factor affecting $r_k(a)$ in an iid multinormal distribution is the distance of the data point a from the mean (d). Since the density in a multinormal distribution decreases with increasing d, $r_k(a)$, an inverse measure of density, will increase with d.

The expected value of $r_k(a)$, i.e., the radius of a hypersphere or ball containing k neighbors in iid multinormal samples was determined by studying randomly generated data for different D, N and k parameters as a function of d, see Fig. 1. These hyperspheres and their radii will be denoted by $H_k^{D,N}(d)$ and $r_k^{D,N}(d)$, respectively. The $r_k^{D,N}(d)$ curves in Fig. 1 were defined for $N=273$ time average circulation patterns (Step 2) using 100 samples of randomly generated multinormal data for each value of $D=13, \ldots, 19$. Each independent variable in the generated data had the same variance as the total variance in the multivariate circulation data (measured by the second moment of the sample.) In each of the 100 experiments, first the appropriate number of multivariate data points ($N=273$) with D independent, normally distributed variables were generated, then for each data point a the distance from all other data points was computed and recorded in increasing order, $r_1^{D,N}(a), \ldots, r_{N-1}^{D,N}(a)$. The radius data were combined for all 100 samples (100×273 measurements) and then grouped into 15–20 classes by d. A spline interpolation based on the 15–20 average values of $r_k^{D,N}(a)$ (not shown in Fig. 1) was used to obtain a continuous estimate of $r_k^{D,N}(d)$, Step 3. The above procedure was repeated for generated data with different D. Similar curves were defined for daily data (changing N and the variance accordingly) and also for subsamples of the time average and daily data sets, using both DGH and RMS distances.

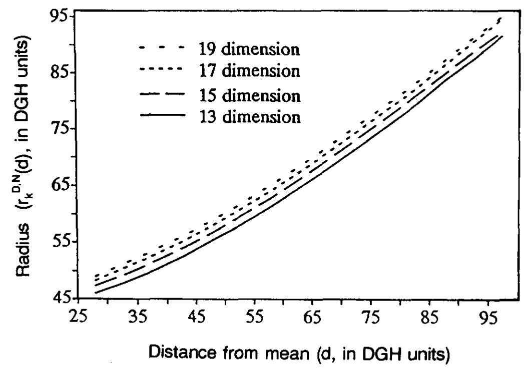

Fig. 1. Radius (in DGH distance) of hyperspheres covering, on average, $k=20$ neighbors out of a sample of $N=273$ data points, as determined for randomly generated multinormal data with different dimensions (D). The average distance of the cases from the theoretical mean of the sample is $d=58.60$ (in DGH units), the same as for the observed time average data set. For further details, see text.

Below, we will measure local density by the number of neighbors (p) contained in the hyperspheres $H_k^{D,N}(d)$. Note that in any phase averaged multinormal data set, whose true dof is $D_t=D$, the expected value of the number of points p ($E(p)$) contained in any $H_k^{D,N}(d)$ is, by definition, k. Similarly if $D_t \neq D$, $E(p) \neq k$. In practice, we can only estimate $E(p)$ through the mean over a sample of points a, $M(p)$. Even if $D_t=D$, $M(p)$ will fluctuate around $E(p)=k$ due to sampling fluctuations. This fluctuation was quantified in the following way. 100 random samples were generated again with a particular D_t. Hyperspheres with radius $r_k^{D,N}(d)$, where $D=D_t$ is the dimension of the generated data, were drawn around each data point a. $M(p)$ was determined as the arithmetic average of p for each sample. Then the $M(p)$ values were arranged in increasing order, $M_1(p), \ldots, M_{100}(p)$. Since $D=D_t$, ${}_{\text{sigl}}p_k^{D,N} = M_5(p)$ and ${}_{\text{sigu}}p_k^{D,N} = M_{96}(p)$ are the lower and upper levels of $M(p)$, respectively, that are statistically significantly different from $E(p)=k$ at, in this example, the 5% significance level (Step 4).

Having determined $r_k^{D,N}(d)$ and ${}_{\text{sigl}}p_k^{D,N}$, ${}_{\text{sigu}}p_k^{D,N}$, the actual test of the dimension of a data set is as follows. We fit $H_k^{D,N}(d)$ hyperspheres around each datapoint in the observed sample. Recall that these hypersheres cover a fixed number of neighbors (k) in iid multinormal data. We will determine the average number of points, $M_k^{D,N}(p)$, covered by the same hypersheres but now in the observed data whose embedding dimension is $n \gg D$ (where n is the number of gridpoints in the data), Step 5. If ${}_{\text{sigl}}p_k^{D,N} < M_k^{D,N}(p) < {}_{\text{sigu}}p_k^{D,N}$, i.e., $M_k^{D,N}(p)$ is close to k, then D is an estimate of the dof in the data set at a particular statistical significance level (Step 6). If $M_k^{D,N}(p)$ is out of the range of the above unequality, we change D and repeat the test. This envolves the use of larger or smaller hyperspheres (see Fig. 1) that cover the same number of neighbors but in an iid multinormal distribution that has a *different* D. The above procedure allows only to determine the intrinsic dimension (Theiler, 1990), since it results, depending on the significance level chosen, in an integer number (or a range of integers) as a dof estimate. This is to be compared to many other methods that seek to determine dimension as accurately as possible. However, as we will see later, the statistical uncertainty due to the limited data precludes a more accurate estimate anyway.

In the next section the dof results will be reported with $k=20$ for the whole daily and time average data sets and with $k=10$ for subsamples of the whole data sets. Experimental estimates (not shown) in the $k=5-20$ hypersphere size range confirmed all the major results to be shown. This is an indication that the self similarity assumption holds well. It means that the structure and dimension of the fractal does not change with decreasing scales in the phase space, at least in this size range (60 m and up for daily data in RMS distance sense.)

3.3. *New aspects of the hypersphere method*

During the past 15 years, surrogate data have been used in several studies to test nonlinearity in univariate time series (see, Theiler et al., 1992). The new aspect of the hyperspheres method described above is that it uses *multivariate* (and not univariate) randomly generated surrogate data. As discussed below, this feature facilitates a direct estimation of dof (or dimension), eliminating the need for the indirect method through scaling, used in all GPM and related studies.

The hyperspheres method, despite the apparent similarities, is different from all GPM type algorithms, including the q-point correlation dimension method (Grassberger, 1983) and the k-nearest-neighbor algorithms (Badii and Politi, 1985). The differences stem from the difference in the assumption about the macro structure of the underlying fractal. Both the correlation and the nearest-neighbor methods assume that the points on the attractor are distributed *uniformly* in the multivariate space and are consequently *unbounded*. In contrast, our assumption is that the distribution is iid *multinormal* and the density strongly depends on the distance from the mean (d).

The assumption of uniformity is unrealistic in most geophysical applications since the variables are clearly confined to a particular range in the phase space. Using a uniform assumption that is not compatible with the observed data leads to problems in the GPM called "edge" effect and limited scaling range (see Subsection 1.2) The net result is an underestimation of dof (and fractal dimension) with GPM. Note that the hyperspheres method overcomes all the problems that are connected with the edge effect and limited scaling range by assuming a bounded distribution and making the hypesphere radii, $r_k^{D,N}(d)$, a function of location (d).

It is important to mention here that the different estimates regarding the minimum number of data points needed for a reliable dimension estimate (Smith, 1988; Ruelle, 1990) are, to a large extent, a reflection of the problems related to the edge effect. Since the hyperspheres method does not suffer from the same problem, these significance estimates are not pertinent to our results, which are without major bias. We note again here that the spectrum of generalized dimensions (D_q) attempts to measure only the extent of lacunarity present in the data. And since the assumption about the uniform macro-structure of the data is unchanged, the edge effect and other problems associated with GPM equally affect the whole spectrum of D_q's.

The other major difference between the hyperspheres method and the GP related methods is in the use of density information. This very important information (along with the sample size N) is basically ignored by the GP and related methods. For this reason, the methods have to involve scaling, i.e., iteratively changing either a distance or mass value in order to arrive at a dof estimate. In contrast, no scaling or iteration is needed in the hyperspheres method which can yield a valid dof estimate with a single application of the algorithm. This is an indication of the power of the current method, which is realized by measuring *density at a given sample size N*. The above differences also explain why much larger number of dof can be reliably estimated by the hyperspheres method, as compared to the GPM.

The hyperspheres method has been tested and its results compared to GPM estimates in simulated data studies with known dof. For example, dof could be ascertained accurately with 24-dimensional iid normal or uniform, randomly generated data (white noise), with only 400 datapoints. GPM estimates, as expected with this small dataset, resulted in serious (20–40%) underestimates.

4. Results

The results of the two statistical methods to assess the dof of the daily and time average, spatially embedded circulation data sets and subsamples of them are summarized in Table 2. Since the method based on the frequency distribution of circulation patterns (hereafter referred to as the χ^2 method) offers not only a way to assess dof but also to test the assumption that the investigated data come from a phase averaged multinormal distribution, results will be first presented from this method.

4.1. χ^2 method

As indicated earlier, only the mean square distance (MSD), a version of RMSD, was applied as a distance function in this method. It can be seen from Table 2 that the assumption of phase

Table 2. *Estimate of independent degrees of freedom of 700 hPa extratropical NH wintertime data sets*

Dataset/ time period	Statistical significance level (%)	Distance function: MSE method: χ^2	MSE hyperspheres	DGH hyperspheres
	Time	Averaged	Data	
all	5	10–14	13–14	16–17
	10	11–14	**14**	16–17
1st half	5	10–12	12–14	16*
2nd half	5	9–24	13–14	17*
		Daily data		
all	5	23–~34	22–24	33–37
	10	24–~33	23–**24**	34–36
1st half	5	NT	21–23	confirms
2nd half	5	NT	22–25	MSE results*
1st 11 years		NT	lowest dimensionality*	
2nd 11 years		NT	higher and approx. equal dimensionality*	
3rd 11 years		NT		

The daily data set contains 330 data points (maps), while the time average data set consists of 273 temporally independent cases. NT indicates cases where no test was performed. An asterisk shows cases where no thorough statistical test was performed. Numbers in bold show the estimate of dimension at the 10% level for the time average and daily data set, using MSE distance function. For further details, see text.

averaged multinormality cannot be rejected for any of the tested data sets at the 5% or even at the 10% significance level. The tests also indicate the range of possible dof values for each data set. For time average data, this is between 11 and 14 at the 10% significance level. For daily data the tests indicate a much higher number of dof, in the range of 24–33. As an example, in Fig. 2 the fit by a standardized χ^2 distribution with 27 dof (the optimal fit in this test) to the empirical frequency values of daily circulation patterns is shown. Note that the fit is quite acceptable even at the tails of the distribution.

We also applied the χ^2 method with *temporal embedding*. Points 10° longitude apart along 50° N were selected individually for this test. The data was sampled and time shifted (delayed) every day. In addition, just as in, e.g., Fraedrich (1986) the seasonal cycle from the data was removed in these experiments. The total number of datapoints was $N = 5000$. For the majority of the geographical locations along 50° N we found that there is a plateau in the dof estimates as n, the embedding dimension, is increased. This plateau is typically around the range of $25 < n < 30$ and is associated with dof estimates between 7 and 13, the most frequent value being 9. The difference between the dof estimates with spatial and temporal embedding will be discussed in Section 5.

4.2. *Hyperspheres method*

The other method, that considers the local density characteristics of the distribution averaged

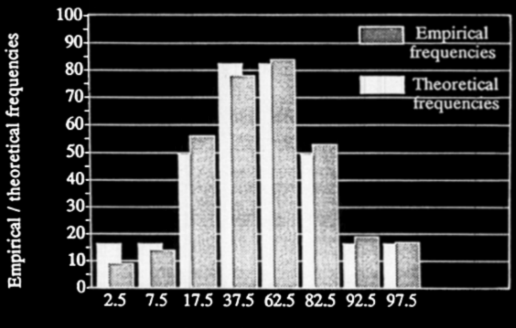

Fig. 2. Empirical frequency distribution of $N = 330$ daily circulation patterns with respect to their distance from the climate mean (d). The classes were defined according to the 5%, 10%, 25%, 50%, 75%, 90% and 95% theoretical expectation from a χ^2 distribution with 27 degrees of freedom. The expected values are also shown.

over all cases (hyperspheres method), was applied with spatial embedding only. In this method, two different distance functions, RMSD and DGH, were used. Let us first examine the results with RMSD, the same distance function as used in the χ^2 test. If both statistical tests are reliable and the underlying attractor is statistically self-similar, at least on the large scales present in the data, we can expect that the range of possible dof obtained by the two methods overlap (and perhaps further narrow the range of dof). As we can see from Table 2 this is the case; the overlap between the results from the two methods also substantiates the credibility of both methods. Looking first at the time average data set, at the 10% level the possible range of dof for the whole sample reduces to one number (14) in the integer domain when the results from the two methods are compared. The same is true for the first 17 winters of the same data set (12 dof) at the 5% level. (Results for the 10% level for subsamples are omitted from Table 2 since they are very close to those at the 5% level.) For the 16 winters from the second half of the time average data set the dof estimate is 13–14 at the 5% (and also at the 10%) level.

As to the daily data set, the comparison of the results from the two methods yields 24 as an estimate for dof at the 10% level. (At the 5% level the estimate would be 23–24). Looking at subsamples of this data set the first half (and especially the first 11 winters) of the data set has fewer dof (21–23) than the rest of the data (22–25). This result is in line with dof estimates for time average data (12 versus 13–14 dof for the first and second half of the data set). Note also that the hyperspheres method, as could be expected from the powerful characteristics of this test (see the second to last paragraph in Section 3), in every case gives a narrower range of estimates for dof in the tested data sets.

When the distance function used in the hyperspheres method is the distance in the gradient of height, higher numbers of dof are detected than when RMSD is applied. When comparing the gradients of the height fields instead of the height fields themselves, the method yields 2–4 extra dof in the time average data set and more than 10 extra dof in the daily data set. This is because successive derivatives of an original pattern possess progressively smaller integral spatial scales.

Another point to note is that the DGH hyperspheres method confirms earlier (RMSD hyperspheres and MSD χ^2) results that the second half of both daily and time average data sets has more dof than the first half. Namias (1980) noted an increased variability in the 1000–700 hPa hemispheric thickness field after 1962. Analyzing 500 hPa hemispheric monthly mean maps Knox et al. (1988) found that a possible climatic "jump" occurred in 1962 after which, among other things, the higher frequency variability increased in the data series. Unfortunately, both studies mentioned above are based on NMC height analysis. Hence the possibility that the statistical results stem from a considerable change in the analysis technique (see, Knox et al., 1988) cannot be ruled out until the data will be reanalyzed (Kalnay et al., 1993.)

5. Discussion

The number of dof in a dataset is not necessarily the dimension of the underlying object. This is because, among other things, the fractal dimension is defined in a limit sense, as the scales in the attractor go to zero. Evidently this requirement cannot be fulfilled with any finite dataset so some extrapolation is necessary to conclude regarding the dimension of an object based on observed data. This is a common problem that all dimension estimation methods and studies face. The extrapolation to infinitesimal scales can be made only if the attractor is self similar, i.e., its behavior and properties do not change as we go from the scales measured to the infinitesimally small scales.

The fact that the global χ^2 method, that uses all scales including the largest in the attractor, and the local hyperspheres method, that uses only the smallest scales available, give overlapping dof estimates suggest that self similarity, at least over the scales present in the data, is not violated seriously. This notion is corroborated by experiments with the hyperspheres method where the scales considered were varied, without any effect on the major results. However, these observations do not give substantial ground for an extrapolation of the results to infinitesimal scales. This is especially true if we consider that there is a very wide range of physical and dynamical processes operating in the atmosphere. At this point one cannot exclude the possibility that at very small scales in the attractor processes that are negligible on the scales that we can measure with our current data sets would become dominant. This would obviously render our dof estimates useless in assessing the fractal dimension of the atmosphere.

Since the large scales in the attractor available to us resolve only the large scale atmospheric circulation, we are inclined to follow Pierrehumbert (1991) and interpret the dof estimates as the dimension of a hypothetical subsystem of the atmosphere that governs the large scale flow. Our general circulation models of the atmosphere could be considered as prototypes of this hypothesized subsystem of the atmosphere. These models do not resolve small-scale motions nor numerous physical processes explicitely. These unresolved processes could actually be considered as miriads of subsystems that are largely independent from each other but interact with the large scale flow. The effect of these unresolved systems is included in large scale models only in a statistical (but usually not stochastic) sense.

Beyond self similarity, there are three more questions we have to raise about the results. The first is about the assumption of iid multinormality. Controlled experiments with the hyperspheres method, where the dof of iid randomly generated data, that is uniformly distributed over an interval (and not normally, as the method assumes), was estimated without significant error, suggests that the exact *shape of the univariate distribution* of variables may not have a substantial impact if the dimension is large. Recall that the GPM assumes an *unbounded uniform distribution* in the multivariate space. The major advantage of the hypersphere method in high dimensions is that the iid normal assumption leads to a *bounded distribution* assumption in the multivariate space; and what univariate distribution we assume may have little impact on the multivariate distribution when dof is high.

The second question we raise is about the quality of data. The geopotential height data used in this study contain noise due to measurement inaccuracy and analysis errors. The size of noise in instantaneous analysis (<20 m in an rms sense, Daley and Mayer, 1986; Kalnay et al., 1993), however, is certainly much below the size of the smallest scales of the attractor studied here (>50 m). Consequently, our dof estimates are not affected directly by the noise present in the data (Ben-Mizrachi et al., 1984).

The third question concerns embedding. Our data set, which is one of the longest series of upper air data, covers only the extratropical NH 700 hPa height level. If we are concerned with the whole atmosphere, then neither our horizontal nor vertical data coverage is complete, not to mention the lack of variables other than height.

Short of complete data coverage, one can only speculate what dimension the global, three dimensional large scale atmospheric circulation might have. Since the current data set spans only a subset of the full attractor, the large-scale atmospheric circulation must have more dimensions than our estimate indicates; the author's guess is somewhere between 60 and 90. This is in the range of Pierrehumbert's (1991) qualitative estimate that has an upper bound of 100. However, even our quantitative estimate for a subsystem, the *spatially embedded* 700 hPa circulation (23–26), is far beyond the range of estimates based purely on *temporal embedding* (3–8).

It has been indicated above that the assumption of self similarity holds for the scales present in our data set. So the large difference between the global (with spatial embedding) and local (with temporal embedding) dof estimates cannot be explained by the difference in the scales these methods consider. The application of the χ^2 method with spatial and temporal embedding resulted in a similarly large difference in dof. This suggests that the likely explanation for resolving the difference must envolve embedding. This conjecture is corroborated by Pierrehumbert's (1991) results. He studied the dof of NH 500 hPa data with a unique combination of spatial and temporal embedding. His GPM dimension estimates are between 20 and 40, clearly in the range of earlier studies using spatial embedding.

To resolve the apparent dichotomy in the results, we follow the lead of Fraedrich (1986) and Tsonis and Elsner (1989) who, in different terms, suggested that the atmosphere is a loosely coupled set of lower-dimensional, regional subsystems. This conceptual framework is also related to the fact that for finite global circulation data, different norms are not equivalent, as shown experimentally for RMS and DGH in this study and theoretically for function spaces (in case of streamfunction and vorticity) in Pierrehumbert (1991).

Lorenz (1991) also pointed out that statistical self similarity may not be true for real life, complex systems. The fine structure of an attractor may indeed be different from its macro structure due to differences in the strength of coupling (or due to the difference among the physical processes operating at various scales.) As he showed, this, in case of temporal embedding, may lead to estimates that reflect the dimension of a subsystem and not the whole system.

According to this interpretation, dynamical dimension estimates based on limited data with *temporal embedding* (in the range of 3–8) will refer to the dimension of an area around the sample location, considered naturally with decreasing weights as the distance from the location increases. These hypothetical regional circulation regimes are not necessarily defined by fixed geographical boundaries. Rather, they manifest themselves by our sampling of the atmosphere at certain locations.

The dimension of hemispheric circulation data, as determined in different studies with spatial embedding to be in the range of 13–44, can be interpreted as a sum of the dimension of such partially overlapping and partially independent (and only loosely coupled) regions or subsystems. Different methods applied to study the spatial correlation characteristics of the circulation (e.g., Wallace and Gutzler, 1981 and Barnston and Livezey, 1987) also reveal that a great portion of the variance in the atmosphere is, indeed, related to regional patterns of a characteristic size of a quarter or third of the hemisphere. To avoid possible misinterpretation of dimension estimates based on temporal embedding, spatial embedding of the phase space (where the scope of the area studied is well controlled by the investigator), where feasible, is also necessary.

6. Conclusions

We presented a new dimension estimation method what we call hyperspheres method. This method focuses on the smallest scales available in a dataset and hence can be called local, just like the Grassberger-Procaccia algorithm and related methods, as opposed to global methods that consider all scales in the attractor. Nevertheless, the hyperspheres method differs from the GPM and related methods in three important aspects. First, the GPM assumes that the data at hand follows a *uniform, unbounded* distribution in the multivariate phase space. Our method, in contrast, assumes that the data follows, in each independent direction, an identical normal distribution. This results in an assumption of a *bounded* distribution in the multivariate space with an expected density strongly dependent on distance from the overall mean of the distribution.

The emphasis here is on the boundedness of the data. Geophysical data is often bounded and can be described vaguely by an iid normal distribution. The hyperspheres method actually attempts to find the dof of the iid normal distribution that can describe the *local density properties* of the sample at a preselected statistical significance level. And here lies the second major difference of the hyperspheres method from the GPM: it uses the essential information of density *explicitly*. As a consequence, no scaling is necessary in the hyperspheres method which can yield a valid estimate with one application of the method. The estimation is actually based on the use of randomly generated multivariate (surrogate) data which constitutes the third important difference from GPM applications.

As a consequence of these differences, the hyperspheres method avoids some problems common to all GPM real data applications. Namely, as was demonstrated in controlled experiments, the hyperspheres method results in accurate dof estimates without a negative bias (that bias is due to the false uniform distribution assumption in GPM), even when dof is high. Furthermore, it avoids all problems related to the infamous "edge effect" (that is due to the unbounded distribution assumption in GPM.) In addition, the problem of limited scaling region (that is also a consequence that GPM assumes a uniform, unbounded distribution) is also solved since the distances (or scales) used in the hyperspheres method depend on distance from the mean of the distribution. And finally, the hyperspheres method provides a clear basis for the evaluation of the statistical significance of the results.

The number of degrees of freedom in 700 hPa Northern Hemisphere circulation data was estimated, beyond the hyperspheres method, by a global (χ^2) method similar to that used by Lorenz (1969). Results from spatial embedding (where NH grid point values serve as embedding variables) was also compared to those with temporal (or time delay) embedding, using the χ^2 method. The possible implications of data quality, statistical assumptions and embedding for the results were also discussed.

The main results of the study are as follows.

(1) It was found that dof estimates from the global and local methods overlap. This indicates that the data set exhibits similar characteristics on the different scales in the attractor, i.e., self similarity is not violated over the scales available in the dataset.

(2) The application of the χ^2 method with spatial and temporal embedding duplicated the large difference in dof found in earlier studies using temporal embedding with local methods (3–8 dofs) and spatial embedding with global methods (13–44 dofs.) This indicates, along with point 1 above, that the large difference between the two groups of studies may be a consequence of difference in embedding and not in the length of scales considered.

(3) The dof estimate of the daily data set (dof = 24) was tentatively interpreted as the dimension of a hypothetical subsystem of the atmosphere that govern the evolution of the large scale flow. Large-scale numerical models of the atmosphere can be considered as examples of such systems.

(4) Dof results with temporal embedding can be interpreted as dimension of geographically confined subsystems within the subsystem of large-scale circulation. These regional systems are only loosely coupled and as far as they can be considered independent from each other, their dimensions can be added up to arrive at the dimension of the large scale global circulation.

(5) The estimate of the dimension of the atmosphere as a whole seems elusive since no ground is evident on which an extrapolation from the relatively large-scales available in observational records could be made to the infinitesimally small phase space scales, required by definition.

7. Acknowledgements

Most of the research was completed in 1989–1990 while the author held a Visiting Fellowship at the Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Colorado. The author is indebted to Maurice Blackmon for his permission to use the computational facilities of the Climate Research Division of the NOAA Climate Monitoring and Diagnostics Laboratory, Boulder, Colorado. Special thanks are due to Donald Mock for his expert help with computation problems. Uwe Radok, Cecil Penland, Henry Diaz and Tamas Prager offered inspiring discussions during the development of this project. The comments of Huug van den Dool, Edward Epstein, Eugenia Kalnay, Ming Cai, Robert Livezey, Peter Caplan, Klaus Fraedrich, Paul Long, James Purser and anonymous reviewers on earlier versions of the manuscript helped to improve its clarity greatly. The manuscript was prepared partly while the author held a National Research Council, NOAA/NMC/ Development Division Research Associateship. The manuscript was typed by Lisa Faunce.

8. Appendix

Eliminating time dependence from a series of circulation patterns

The idea behind the statistical test applied to check the time dependence present in the series of 290 time average maps (see Subsection 2.1) is the following. If there is no time dependence in the series then the average of the one time step distances between maps should be approximately the same as the average distance between maps selected *randomly* from the same period (± 45 days) of the year. So first the average of one time step distances was calculated for the whole data series excluding the lags from the end of one winter season to the beginning of the next (average of 256 RMSD or DGH values). A similar average was computed for the same 256 cases but now using distances between each circulation map and a randomly chosen map from any other year (but from the same season), instead of using distances between successive maps. This random experiment was repeated 100 times, resulting in a series of random average distance values. Since the average distance for one time step lags for the original data series ranked second in the ordered series of 100 average random distances, the null hypothesis of time independence was rejected at the 5% level.

To eliminate time dependence we removed some maps from the original series. This was done, after experimenting with the parameters of the method, in the following way. For each circulation pattern in the data series we determined the 5 closest patterns in the rest of the data set. If the successive circulation pattern was found among the 5 closest patterns, it was considered a redundant element in the data series and consequently was removed. Altogether 17 circulation patterns were thus removed. The average one-time-step distance for the resulting 273 element data series ranked as 135th in the ordered series of 1000 random average distances. Therefore the null hypothesis of time independence for this reduced time series was not rejected at the 10% (or any lower) significance level. The 273 time average circulation maps were then used to determine and test the number of degrees of freedom in the data.

REFERENCES

Badii, R. and Politi, A. 1985. Statistical description of chaotic attractors: the dimension function. *J. Stat. Phys.* **40**, 725–750.

Barnett, T. P. and Preisendorfer, R. 1987. Origins and levels of monthly and seasonal forecast skill for United States surface air temperatures as determined by canonical correlation analysis. *Mon. Wea. Rev.* **115**, 1825–1850.

Barnston, A. G. and Livezey, R. E. 1987. Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Mon. Wea. Rev.* **115**, 1083–1126.

Ben-Mizrachi, A., Procaccia, I. and Grassberger, P. 1984. Characterization of experimental (noisy) strange attractors. *Physical Review* **A29**, 975–977.

Daley, R. and Mayer, T. 1986. Estimate of global analysis error from the Global Weather Experiment Observational Network. *Mon. Wea. Rev.* **114**, 1642–1653.

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

Fraedrich, K. 1986. Estimating the dimensions of weather and climate attractors. *J. Atmos. Sci.* **43**, 419–432.

Fraedrich, K. 1987. Estimating weather and climate predictability on attractors. *J. Atmos. Sci.* **44**, 722–728.

Fraedrich, K., Ziehmann, Ch. and Sielmann, F. 1995. Estimates of spatial degrees of freedom. *J. Climate* **8**, 361–369.

Grassberger, P. 1983. Generalized dimensions of strange attractors. *Phys. Lett.* **A97**, 227–230.

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

Halsey, T. C., Jensen, M. H., Kadanoff, L. P., Procaccia, I. and Shraiman, B. I. 1986. Fractal measures and their singularities: the characterization of strange sets. *Phys. Rev.* **A33**, 1141–1151.

Henderson, H. W. and Wells, R. 1988. Obtaining attractor dimensions from meteorological time series. *Advances in Geophysics* **30**, 205–237.

Hentschel, H. G. E. and Procaccia, I. 1983. The infinite number of generalized dimensions of fractals and strange attractors. *Physica* **8D**, 435–444.

Horel, J. D. 1985. Persistence of the 500 mb height field during Northern Hemisphere winter. *Mon. Wea. Rev.* **113**, 2030–2042.

Kendall, M. G. and Stuart, A. 1967. The advanced theory of statistics, vol. 2. Hafner Publishing Co., New York, 420–441.

Keppenne, C. L. and Nicolis, C. 1989. Global properties and local structure of the weather attractor over western Europe. *J. Atmos. Sci.* **46**, 2356–2370.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Saha, S., White, G., Woolen, J., Chelliah, M., Janowiak, J., Mo, K. C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., Kung, E. and Salstein, D. 1993. *The NMC/NCAR CDAS/reanalysis project.* NMC Office Note 401. Available: NOAA/NMC, World Weather Building, 5200 Auth Rd., Camp Springs, MD 20746.

Knox, J. L., Higuchi, K., Shabbar, A. and Sargent, N. E. 1988. Secular variation of Northern Hemisphere 500 hPa geopotential height. *J. Climate* **1**, 500–511.

Livezey, R. E. and Chen, W. Y. 1983. Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.* **111**, 49–59.

Lorenz, E. N. 1969. Atmospheric predictability by naturally occurring analogs. *J. Atmos. Sci.* **26**, 636–646.

Lorenz, E. N. 1991. Dimension of weather and climate attractors. *Nature* **353**, 241–244.

Namias, J. 1980. Some concomitant regional anomalies associated with hemispherically averaged temperature variations. *J. Geophys. Res.* **85**, 1585–1590.

O'Lenic, E. A. and Livezey, R. E. 1988. Practical considerations in the use of Rotated Principal Component Analysis (RPCA) in diagnostic studies of upper-air height fields. *Mon. Wea. Rev.* **116**, 1682–1689.

Panofsky, H. A. and Brier, G. W. 1968. *Some applications of statistics to meteorology.* Pennsylvania State University Press, 224 pp.

Pierrehumbert, R. T. 1991. Dimensions of atmospheric variability. In: *Beyond belief: randomness, prediction and explanation in science*, eds: J. L. Casti and A. Karlqvist. CRC Press, pp. 110–142.

Ruelle, D. 1990. Deterministic chaos: the science and fiction. *Proc. R. Soc. Lond.* **A427**, 241–248.

Smith, L. A. 1988. Intrinsic limits on dimension calculations. *Physics letters* **A133**, 283–288.

Takens, F. 1981. Detecting strange attractors in turbulence. *Dynamical systems and turbulence.* Warwick; Springer-Verlag, 336–381.

Theiler, J. 1990. Estimating fractal dimension. *J. Opt. Soc. Am.* **A7**, 1055–1073.

Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J. D. 1992. Testing for nonlinearity in time series: the method of surrogate data. *Physica* **D58**, 77–94.

Toth, Z. 1991a. Circulation patterns in phase space: A multinormal distribution? *Mon. Wea. Rev.* **119**, 1501–1511.

Toth, Z. 1991b. Intercomparison of circulation similarity measures. *Mon. Wea. Rev.* **119**, 55–64.

Toth, Z. 1992. Quasi-stationary and transition periods in the NH circulation series. *J. Climate* **5**, 1235–1247.

Toth, Z. 1993. Preferred and unpreferred circulation types in the Northern Hemisphere wintertime phase space. *J. Atmos. Sci.* **50**, 2868–2888.

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

Van den Dool, H. M. and Chervin, R. M. 1986. A comparison of month-to-month persistence of anomalies in a general circulation model and in the Earth's atmosphere. *J. Atmos. Sci.* **43**, 1454–1466.

Vautard, R. and Ghil, M. 1989. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimate time series. *Physica* **35D**, 395–424.

Wallace, J. M. and Gutzler, D. S. 1981. Teleconnections in the geopotential height field during the NH winter. *Mon. Wea. Rev.* **109**, 784–812.