ELSEVIER

19 September 1994

PHYSICS LETTERS A

Physics Letters A 193 (1994) 67–74

Testing for nonlinearity in weather records

Milan Paluš [a,1], Dagmar Novotná [b]

[a] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA*

[b] *Institute of Atmospheric Physics, Academy of Sciences of the Czech Republic, Boční II/1401, 141 31 Prague 4, Czech Republic*

Received 24 May 1994; revised manuscript received 22 July 1994; accepted for publication 26 July 1994
Communicated by A.R. Bishop

Abstract

Daily records of atmospheric surface pressure, temperature and geopotential heights of 500 hPa isobaric level were tested for nonlinearity, the necessary condition for deterministic chaos, using redundancy and surrogate data techniques. While the time series of the temperature and the geopotential heights were found indiscernible to be from correspondent isospectral linear stochastic processes, a significant nonlinear component was detected in the dynamics of the pressure recording, however, no specific signatures of low-dimensional chaos were manifest.

During the last decade many papers devoted to the problem of inferring the dynamical mechanisms of the weather and climate changes from recorded data have been published. The measured quantities selected for the analyses, have included, e.g., local surface pressures, relative sunshine durations, zonal wave amplitudes [1], upper-level geopotential heights [2,3], low-level vertical velocity components [4], or, oxygen-isotope concentrations in deep sea cores [1,5–8]. In the majority of the cases the Grassberger–Procaccia algorithm for estimating the correlation dimension [9,10] was used as the analytical tool, and low values of the dimension estimates obtained were claimed as evidence for low-dimensional chaos in the weather or climate dynamics [1–5,7]. On the other hand, Grassberger [6] cautioned that in the case of short and noisy data, as the climatic and weather records usually are, the reliability of the method is questionable and the low values of the dimension estimates may be spurious. And indeed, he constructed a random series of corresponding length, preprocessed by the same way as the climatic record in Ref. [5] and obtained a low value of the estimated dimension. Also Lorenz [11] writes that it seems unlikely that global weather or climate systems possess a low-dimensional attractor.

The problem of reliability of dimensional or Lyapunov exponent algorithms, applied to experimental data, is the general problem of nonlinear time-series analysis, and spurious results, leading to false identification of chaotic dynamics in data, consistent with a simpler explanation, can emerge elsewhere [12–16]. Some authors, considering these problems, proposed to test necessary conditions for chaos, like nonlinearity or nonlinear determinism [17–19]. As was pointed out by Theiler et al. [17], detection of nonlinearity is a considerably easier goal than that of positively identifying chaotic dynamics. On the other hand, when the scrutinized data is found consistent with the explanation by a linear stochastic process, underlying low-dimensional chaos is improbable.

Recently, Paluš et al. [18] have demonstrated how the information theoretic functionals – redundancies

[1] E-mail: mp@santafe edu

0375-9601/94/$07.00 © 1994 Elsevier Science B.V. All rights reserved
SSDI 0375-9601(94)00639-3

– can be used for the detection of nonlinearity. A brief overview of the method follows

Let $X_1, X_2, \ldots, X_n$ be an n-dimensional random variable with zero mean, covariance matrix $\mathbf{C}$ and probability distribution density (PDD) $p(x_1, x_2, \ldots, x_n)$ The individual PDDs are denoted as $p(x_i)$, $i = 1, \ldots, n$ (For convenience, we use the notation $p(x_i)$, rather than the more accurate $p_{X_i}(x_i)$, and an analogous simplification for the n-dimensional PDD)

Then the *redundancy* $R(X_1, \ldots, X_n)$, measuring the level of mutual dependence of the components $X_1, X_2, \ldots, X_n$, is

$$R(X_1, \ldots; X_n) = \int \ldots \int p(x_1, x_2, \ldots, x_n) \times \log \frac{p(x_1, x_2, \ldots, x_n)}{p(x_1)p(x_2)\ldots p(x_n)} \mathrm{d}x_1 \ldots \mathrm{d}x_n \quad (1)$$

The redundancy (1) for $n = 2$ is called mutual information. For relations of redundancies and entropies, and for further details see Refs [20,21,18,23] and references therein.

Further, we define the *linear redundancy* $L(X_1, \ldots, X_n)$ of $X_1, X_2, \ldots, X_n$ as

$$L(X_1; \ldots, X_n) = \tfrac{1}{2} \sum_{i=1}^{n} \log(c_{ii}) - \tfrac{1}{2} \sum_{i=1}^{n} \log(\sigma_i), \quad (2)$$

where c_{ii} are the diagonal elements (variances) and σ_i are the eigenvalues of the $n \times n$ covariance matrix $\mathbf{C}$

If $X_1, \ldots, X_n$ has an n-dimensional Gaussian distribution, then $L(X_1, \ldots, X_n)$ and $R(X_1, \ldots, X_n)$ are theoretically equivalent [24]

In usual experimental situations one records a time series $\{y(t)\}$ of a specific observable. $\{y(t)\}$ is usually considered as a realization of a stochastic process $\{Y(t)\}$, which is stationary and ergodic

We will study the redundancies for the variables

$$X_i(t) = y(t + (i-1)\tau), \quad i = 1, \ldots, n, \quad (3)$$

where τ is a time delay and n is an embedding dimension [22] Redundancies of the type

$$R(y(t); y(t+\tau), \ldots, y(t + (n-1)\tau))$$

are, due to stationarity of $\{Y(t)\}$, independent of t. We introduce the notation

$$R^n(\tau) = R(y(t), y(t+\tau), \ldots, y(t + (n-1)\tau)) \quad (4)$$

for the redundancy and

$$L^n(\tau) = L(y(t); y(t+\tau), \ldots, y(t + (n-1)\tau)) \quad (5)$$

for the linear redundancy of the n variables $y(t)$, $y(t+\tau)$, $\ldots$, $y(t + (n-1)\tau)$ (Quantities (4) and (5) are obtained from a single process realization – an experimental time series – by time averaging, which can be applied due to the above assumption of ergodicity)

The redundancy $R^n(\tau)$, based on PDDs, measures general dependences among the series $\{y(t)\}$ and its delayed versions, while the linear redundancy $L^n(\tau)$, based on covariances (or correlations, see Refs. [18,23]) reflects only their linear relations Therefore Paluš et al [18] proposed to compare $L^n(\tau)$ with $R^n(\tau)$, considered as the functions of the time lag τ If their shapes are the same or very similar, a linear description of the process under study should be considered sufficient. Large discrepancies suggest important nonlinearities in the dynamics of the process under study.

This approach, further referred to as *qualitative testing*, or, *qualitative comparison*, was demonstrated to be successful in distinguishing specifically nonlinear dynamics (chaotic [18] or also nonchaotic [23]) from linear stochastic processes. In comparison of figures, however, there can be a subjective factor, influencing the conclusion about the findings. Therefore the development of a quantitative method, suitable for rigorous statistical testing, is desirable This can be achieved by incorporating the concept of "surrogate data", advocated by Theiler et al [17]

In testing for nonlinearity, the surrogate data are artificially generated realizations of a linear stochastic process, which mimic linear properties of the studied data, namely the spectrum and autocorrelation function (For details see Refs. [17,25]) Nonlinear structures, if present in the scrutinized data, are eliminated in the surrogates. Thus we can statistically compare the redundancies obtained from the analyzed data and from a set of surrogates to find out whether the data is significantly different from the null hypothesis of a linear stochastic process, or whether a linear stochastic process is a probable explanation of the data dynamics. As the discriminating statistic we use the difference between the redundancy, obtained from the data, and the mean value of the redundancy for a set of surrogates, in the number of standard deviations

(SDs) of the latter. The null hypothesis is rejected and the results called *significant*, when the probability p of the null hypothesis is lower than a chosen level, usually set to 0 05 or 0.01. Providing the statistic has a normal distribution $N(0,1)$ and 30 realizations of the surrogates are used in the test [2], then the probability of the null hypothesis is $p < 0.05$ for values of the statistic greater than 1 699; or greater than 2 462 for $p < 0.01$, etc., see e.g. Ref. [26]

It is well-known that in construction of n-dimensional embeddings by the time delay method, the results are usually influenced by the choice of the embedding dimension n and the time delay τ. Therefore, in the quantitative testing, we evaluate the redundancy-based statistics for a wide range of time delays and, if possible (i e., allowed by the time series length), also for several embedding dimensions.

However, one cannot always construct good surrogate data. That is, in spite of theoretical expectation, in numerical practice linear properties of the surrogates can differ from those of the data under study. The changes in linear properties are reflected in nonlinear measures [3] as well, and thus, a change in linear properties can lead to spurious detection of nonlinearity in linear data [27,25]. Therefore, we evaluate also the statistic based on the linear redundancy $L^n(\tau)$, reflecting specifically the changes in linear properties Then only those significant differences in the nonlinear statistic can reliably count for nonlinearity, which are not detected in the linear statistic.

Due to the central limit theorem, the distribution of the isospectral surrogates of a stationary process (generated by the Fourier transform) tends to a Gaussian distribution. If a (one-dimensional) distribution of the scrutinized data is different from the Gaussian one, another case of spurious identification of nonlinearity can emerge. Therefore, before the analysis, we perform a histogram transformation – "Gaussianization", resulting in an approximate Gaussian distribution of the data [25,28].

Data Two series of daily values of geopotential heights of 500 hPa isobaric level were analyzed, the first, 6570 samples (18 years) recorded in Prague, Ruzyně station, the second, 11670 samples (32 years) recorded in Krakow. We generated the surrogate data using the fast Fourier transform (FFT), which requires the series length to be a power of two, therefore we analyzed subseries of lengths 4096 and 8192 samples, respectively

The other two analyzed series, recorded in Prague, Klementinum station, are more unique· the series of 200 years (73000 samples) of mean daily values of the surface atmospheric temperature and daily values of the surface atmospheric pressure. Again, due to FFT-based surrogate, we analyzed subseries of 65536 samples

Results Geopotential heights The results of the analysis of the Prague series of the geopotential heights are presented in Fig. 1. The qualitative comparison shows no substantial difference between the time plots of the linear redundancy $L^2(\tau)$ and the redundancy $R^2(\tau)$, showed in Figs 1a and 1b, respectively In the quantitative analysis (Figs 1c, 1d) there are several formally significant results (i e differences greater than 1 699), however, there are two reasons why not to reject the null hypothesis of a linear stochastic process·

(a) Statistical reason. Due to multiplicity of the test values (60 in this case) the criterion for significance of an individual value must be strengthened, i.e., based on the Bonferroni inequality we should take, in this case, $p < 0.05/60$ instead of $p < 0.05$ [29,30], which increases the critical value of the statistics from 1.699 to approximately 3.5 This approach, however, is fully correct for independent test values, for the dependent test values, which is the case here, the power of the test can be decreased. In order to avoid the type II error (i.e., acceptance of the null hypothesis when it should be rejected), the Heilperin–Ruger inequality can be considered instead of the Bonferroni inequality, and, expecting k significant values (from m total test values) $p < 0.05k/m$ can be taken [31,32]. In this case the critical value is still about 3. Thus, no significant difference was found.

(b) Methodological reason: Even if we accept some

---

[2] With a limited number N_S of the realizations of the surrogates, the t-distribution with $N_S - 1$ degrees of freedom should be used for deriving the critical values of statistics, instead of the normal distribution In all the tests, presented here, 30 realizations of surrogates were used, therefore the critical values presented in this paper are related to the t-distribution with 29 degrees of freedom and to one-sided tests, because the change of the statistics are expected in one direction For more details see Ref [25]

[3] We use primarily $R^n(\tau)$, other authors use dimensions, correlation integral, or nonlinear forecastibility [17]

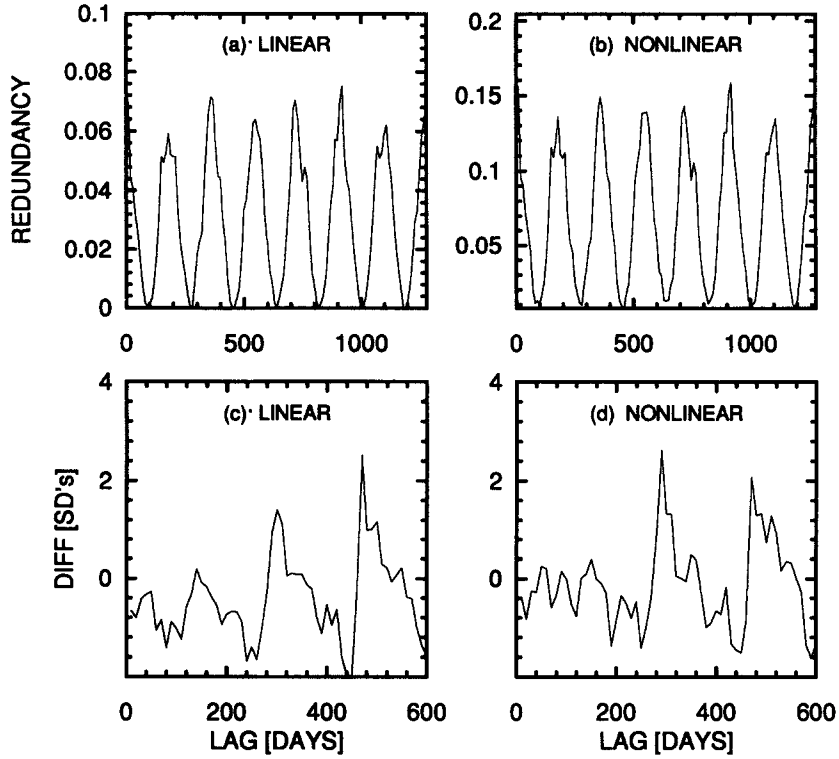

Fig. 1. (a) Linear redundancy $L^2(\tau)$, (b) redundancy $R^2(\tau)$, as functions of the lag τ, (c) linear redundancy statistic, (d) redundancy statistic for the Prague series of the geopotential heights of the 500 hPa isobaric level. Embedding dimension $n = 2$

values of the nonlinear statistic (based on $R^n(\tau)$) as significant, we cannot reject reliably the null hypothesis of a linear stochastic process, as far as equivalent differences were found in the linear statistic (based on $L^n(\tau)$). Therefore the observed differences can be caused by the fact that the surrogates do not exactly mimic the linear properties of the data, not by a nonlinearity.

We can conclude that by both the qualitative and quantitative methods we found the Prague series of the geopotential heights indiscernible from the isospectral linear stochastic process

The main feature of the dynamics of the above geopotential heights series, as it can be observed in the time-lag plots of the redundancies (Figs. 1a, 1b), is the one-year periodicity. We can ask whether there is anything beyond this dynamics; therefore we analyzed also the filtered series, in which one-year periodicity was eliminated by the FFT based filter. In the qualitative analysis both $L^n(\tau)$ and $R^n(\tau)$ show the same picture – they decrease quickly until the lag 12 days and then fluctuate about a very low level. The question whether these small values ($(2–4)\times10^{-3}$) can mean a "numerical zero", i.e., the fact that the filtered series $\{y(t)\}$ and $\{y(t+\tau)\}$ for $\tau > 12$ are independent, was answered by the quantitative test using so-called scrambled surrogates – the elements of the series were mixed in temporal order so that all temporal correlations were destroyed. Comparing the data with the scrambled surrogates the null hypothesis of an IID (independent identically distributed) process was tested and rejected (differences of 4–8 SDs) On the other hand, using the FFT surrogates, both the stronger dependences (the lags 1–12) and the weak dependences for the lags $\tau > 12$ days were found consistent with the isospectral linear stochastic process.

The results for the Krakow series were very similar

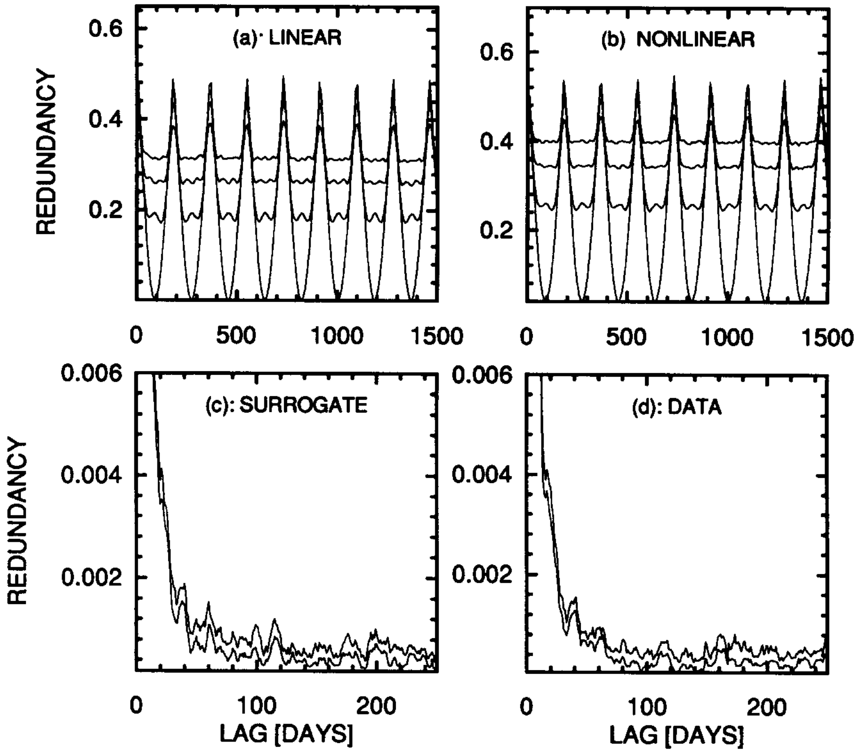

Fig. 2 (a) Linear redundancy $L^n(\tau)$, (b) redundancy $R^n(\tau)$, computed from the surface atmospheric temperature series. The four different curves in each picture are the redundancies for different embedding dimensions, $n = 2$–5, reading from the bottom to the top. Redundancies $L^n(\tau)$ and $R^n(\tau)$ are plotted as $L^n(\tau)/(n-1)$ and $R^n(\tau)/(n-1)$, respectively. (c) $R^n(\tau)$, $n = 2$ (lower curve), 3 (upper curve), for the surrogate data of the filtered temperature series, (d) $R^n(\tau)$, $n = 2$ (lower curve), 3 (upper curve), for the filtered temperature series.

to those for the above Prague series. Therefore, we can conclude that the analysis of the recordings of the geopotential heights did not yield any argument to reject the linear stochastic explanation.

Temperature The results of the analysis of the surface atmospheric temperature record (65536 samples - days) are presented in Fig. 2. The qualitative analysis of the data (Figs. 2a, 2b) brought no substantial difference between $R^n(\tau)$ and $L^n(\tau)$, $n = 2$–5, $\tau = 10$–1500 days (Figs 2a, 2b). (Analyzed, but not presented, were also short-time dependences for $\tau = 1$–250 days.) In the quantitative analysis, the differences obtained were not higher than 1.6 SDs. After filtering out the one-year periodicity, the quantitative analysis brought no significant results, like the analysis before the filtration. In the qualitative analysis (Figs. 2c, 2d: we present $R^n(\tau)$ computed from the data and its surrogate, this comparison is equivalent to those of $R^n(\tau)$ and $L^n(\tau)$ from the data [18,23,25]) we can see, that the redundancies of the filtered temperature series decrease until the lag of about 80 days and then fluctuate about the same (low) level. Similarly, like in the case of the geopotential heights data, the hypothesis of an IID process was rejected, however, all those dependences were found consistent with a linear stochastic explanation.

The above analysis of the temperature record brought no arguments to reject the null hypothesis of a linear stochastic process.

Pressure The results of the analysis of the surface atmospheric pressure record (65536 samples - days) are presented in Fig. 3. The qualitative analysis of the data (Figs. 3a, 3b) shows some differences between $L^n(\tau)$ and $R^n(\tau)$, namely the half-year peaks are not

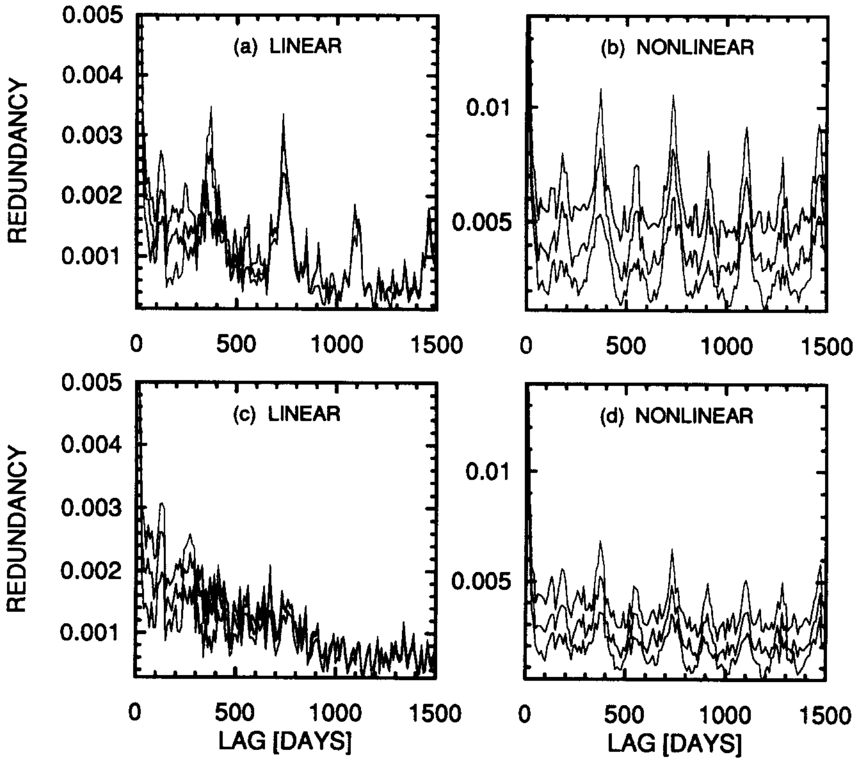

Fig. 3. (a) Linear redundancy $L^n(\tau)$, (b) redundancy $R^n(\tau)$, computed from the surface atmospheric pressure series. (c) Linear redundancy $L^n(\tau)$, (d) redundancy $R^n(\tau)$, computed from the filtered surface atmospheric pressure series. The three different curves in each picture are the redundancies for different embedding dimensions, $n = 2$–4, reading from the bottom to the top. Redundancies $L^n(\tau)$ and $R^n(\tau)$ are plotted as $L^n(\tau)/(n-1)$ and $R^n(\tau)/(n-1)$, respectively.

so clearly pronounced in $L^n(\tau)$ as in $R^n(\tau)$, the one-year periodicity, however, is apparent in both $L^n(\tau)$ and $R^n(\tau)$. Clearer results were obtained by the quantitative analysis: while no significant differences were detected on the linear level, i.e., by the statistic based on $L^n(\tau)$, the nonlinear statistic (based on $R^n(\tau)$) brought significant differences of values between 5 and 15 SDs. The results of the analysis of the filtered pressure series are even more illustrative. The results of the quantitative analysis did not change after the filtration, i.e., evidence for nonlinearity, safe from spurious effects of differences on the linear level, was detected. In the qualitative comparison (Figs. 3c, 3d), we can see that $R^n(\tau)$ decreased after the filtration, i.e., the linear contribution to the dependence structures in the data (reflected in nonlinear $R^n(\tau)$ as well) was removed by the filtration, while the character of the time-lag dependence of $R^n(\tau)$ is almost the same as in $R^n(\tau)$ computed from the original data, i.e., principal one-year peaks and smaller half-year peaks were detected. On the other hand, linear redundancy $L^n(\tau)$ of the filtered data does not reveal these structures. The latter is also evidence that the filtration was well done. If the periodicity was due to a peak in the spectrum, it would be detected by the linear redundancy as well. (An example of a numerically generated nonlinear series, in which a periodic structure was detected only by the redundancy $R^n(\tau)$ and not by the linear redundancy $L^n(\tau)$, can be found in Refs. [23,25].)

Thus, we can conclude, that the results of both the qualitative and quantitative methods show that a linear stochastic explanation of the pressure series is not adequate and the data contains an important nonlinear component.

The indiscernibility of the geopotential heights and the temperature series from linear stochastic processes, demonstrated above, supposes that the detections of low dimensionality in similar studies were probably spurious On the other hand, our results cannot be understood as evidence that the nature of corresponding atmospheric processes is indeed linear stochastic. Even if the dynamics of a particular process, or a particular atmospheric subsystem, is nonlinear and deterministic, as many physical models propose, its numerous interactions with other subsystems and disturbances of various origins and scales, can influence a particular measurement in such a way, that the best explanation of particular data, based on its analysis, is as a linear stochastic process. Simulation experiments, involving networks of coupled dynamical systems, related to the models studied in the atmospheric physics, could be of interest here.

In the case of the pressure series, an important nonlinear component was detected. In our estimates of the correlation dimension, however, no saturation was observed up to embedding dimension 12 Also, no specific features of low-dimensional chaos were observed in the analysis based on the marginal redundancy technique [23]. Thus we can conclude that the daily recording of the surface atmospheric pressure has nonlinear dynamics, but it is not a case of low-dimensional chaos. An explanation by a high-dimensional, deterministic, nonlinear, and maybe chaotic process is possible, however, the hypothesis of a nonlinear stochastic process is, at least formally, equivalent, due to the practical impossibility to distinguish the two by the recent method of nonlinear time series analysis.

MP would like to thank Dean Prichard for inspiring discussions and critical reading of the manuscript. Helpful discussions with J Theiler, J. Klaschka and V. Witkovský are also gratefully acknowledged MP was supported by the International Research Fellowship F05 TWO4757 ICP from the National Institutes of Health, Fogarty International Center, and also by grants to the Santa Fe Institute, including core funding from the John D. and Catherine T MacArthur Foundation, the National Science Foundation (PHY-8714918), and the U.S Department of Energy (ER-FG05-88ER25054) DN was supported by grant No. 34209 from the Academy of Sciences of the Czech Republic.

References

[1] K Fraedrich, J. Atmos Sci 43 (1986) 419
[2] C Essex, T. Lookman and M.A.H Neremberg, Nature 326 (1987) 64
[3] C L Keppenne and C J Nicolis, J Atmos Sci 46 (1989) 2356
[4] A A Tsonis and J B Elsner, Nature 333 (1988) 545
[5] C Nicolis and G Nicolis, Nature 311 (1984) 529
[6] P Grassberger, Nature 323 (1986) 609
[7] C Nicolis and G Nicolis, Nature 326 (1987) 523
[8] P Grassberger, Nature 326 (1987) 524
[9] P Grassberger and I Procaccia, Phys Rev Lett 50 (1983) 346
[10] P Grassberger and I Procaccia, Physica D 9 (1983) 189
[11] E N Lorenz, Nature 353 (1991) 241
[12] G Mayer-Kress, in Directions in chaos, ed Hao Bailin (World Scientific, Singapore, 1987) p. 122
[13] P E Rapp, A M Albano, T I Schmah and L A Farwell, Phys Rev E 47 (1993) 2289
[14] A R Osborne and A Provenzale, Physica D 35 (1989) 357
[15] L A Smith, Phys Lett A 133 (1988) 283
[16] J Theiler, Phys Rev A 34 (1986) 2427
[17] J Theiler, S Eubank, A Longtin, B Galdrikian and J D. Farmer, Physica D 58 (1992) 77
[18] M Paluš, V Albrecht and I Dvořák, Phys Lett A 175 (1993) 203
[19] D T Kaplan and L Glass, Phys Rev Lett 68 (1992) 427
[20] T M Cover and J A Thomas, Elements of information theory (Wiley, New York, 1991)
[21] A M Fraser, IEEE Trans Inf Theory 35 (1989) 245
[22] F Takens, in Lecture notes in mathematics, Vol 898 Dynamical systems and turbulence, Warwick, 1980, eds D A. Rand and D S Young (Springer, Berlin, 1981) p 365
[23] M Paluš, in Santa Fe Institute Studies in the Sciences of Complexity, Proc Vol XV Time series prediction forecasting the future and understanding the past, eds A S Weigend and N A Gershenfeld (Addison-Wesley, Reading, MA, 1993) p 387
[24] S D Morgera, IEEE Trans Syst Man Cybern SMC-15 (1985) 608
[25] M Paluš, Testing for nonlinearity using redundancies quantitative and qualitative aspects, Santa Fe Institute Working Paper 93-12-076, Physica D, in press
[26] P J Bickel and K A Doksum, Mathematical statistics (Prentice Hall, Englewood Cliffs, NJ, 1977)
[27] J Theiler, P S Linsay and D M Rubin, in Santa Fe Institute Studies in the Sciences of Complexity, Proc Vol. XV Time series prediction forecasting the future and understanding the past, eds A S Weigend and N A Gershenfeld (Addison-Wesley, Reading, MA, 1993) p 429
[28] M B Kennel and S Isabelle, Phys Rev A 46 (1992) 3111

[29] R.G. Miller, Simultaneous statistical inference (Springer, Berlin, 1980).
[30] F Mosteller and J.W Tukey, Data analysis and regression (Addison-Wesley, Reading, MA, 1977)
[31] G Hommel, Biometrical J. 25 (1983) 423
[32] D Morgenstern, Metrika 27 (1980) 171
[33] K Fraedrich, J Atmos. Sci 44 (1987) 722