Determination of Kolmogorov Entropy of Chaotic Attractor Included in One–Dimensional Time Series of Meteorological Data

Yan Shaojin (严绍瑾), *Peng Yongqing* (彭永清)
and *Wang Jianzhong* (王建中)
Nanjing Institute of Meteorology, Nanjing 210044, China
Received June 5, 1990; revised September 18, 1990

ABSTRACT

The 1970–1985 day to day averaged pressure dataset of Shanghai and the extension method in phase space are used to calculate the correlation dimension D and the second–order Renyi entropy K_2 of the approximation of Kolmogorov's entropy, the fractional dimension $D = 7.7 \sim 7.9$ and the positive value $K_2 \approx 0.1$ are obtained. This shows that the attractor for the short–term weather evolution in the monsoon region of China exhibits a chaotic motion. The estimate of K_2 yields a predictable time scale of about ten days. This result is in agreement with that obtained earlier by the dynamic–statistical approach.

The effects of the lag time τ on the estimate of D and K_2 are investigated. The results show that D and K_2 are convergent with respect to τ. The day to day averaged pressure series used in this paper are treated for the extensive phase space with $\tau = 5$, the coordinate components are independent of each other; therefore, the dynamical character quantities of the system are stable and reliable.

I. INTRODUCTION

The Kolmogorov entropy or metric entropy (hereafter designated K–entropy) is an essential physical quantity in the study of the evolution of a nonlinear system. The value of K–entropy can be used to distinguish regular motions from irregular ones, i.e., the motion is ordered when $K = 0$, stochastic when $K \to \infty$, and chaotic when $K > 0$.

It is indicated in modern research that the fractional dimension D, the Lyapunov exponent λ_i and the K–entropy of a dynamic system are physical quahtities to judge the characters of the system in phase space, they are characteristic quantities for showing whether the motion is chaotic or not. Undoubtedly they are closely interrelated. Recent studies show that the value of K–entropy is able not only to describe the nature of the motion in phase space, and more interesting thing is that its reciprocal, $T = 1 / K$, is frequently used to estimate the mean predictable time scale of the system; therefore, the research of K–entropy is attracting more and more scientists including meteorologists.

We, using the method of extending phase space, have shown that there exist attractors with $d_s = 3.4$ (Shanghai) and $d_g = 2.3$ (Guangzhou) for the short–term climate system, based upon the data of 1873–1980 monthly mean temperatures of the two cities (Peng et al., 1989), later we calculated (Yan et al., 1990) the lyapunov exponents and obtained $\lambda_{1,2} > 0$ and $\lambda_3 < 0$, using the Wolf's method (Wolf et al., 1985), for the 1873–1980 Guangzhou monthly mean temperatures. All these results demonstrated that the short–range climate evolution in monsoon area of China is a kind of chaotic motion.

In this paper, we draw on the data of 1970–1985 day to day averaged pressures of Shanghai to investigate the short–term weather evolutions in monsoon area and to emphasize the function of climate attractor. The basic idea is through the discussion of correlation dimension and Lyapunov exponent, especially the calculation of K–entropy, to estimate the predictable time scale of short term climate system.

II. DETERMINATION OF K–ENTROPY AND CALCULATION SCHEME

Consider a dynamic system with N degrees of freedom, as a measurement the N–dimensional phase space can be divided into many boxes with volume, ε^N, ε is a resolution scale of measurement. Assume that in the phase space exists an attractor on which is the trajectory $\vec{x}(t)$. We measure the state of the attractor at time interval k. If the trajectory $\vec{x}(t=k)$ is in box i_1, $\vec{x}(t=2k)$ in box $i_2, \cdots\cdots \vec{x}(t=dk)$ in box i_d, then the joint probability for this kind of distribution is $P(i_1, i_2, \cdots, i_d)$ and K–entropy is defined as

$$K = -\lim_{k\to 0}\lim_{\varepsilon\to 0}\lim_{d\to\infty}\frac{1}{dk}\sum_{i_1\cdots i_d} p(i_1,\cdots,i_d)\ln p(i_1,\cdots,i_d) \tag{1}$$

This formula is similar to that of information entropy defined by Shannon. In fact, it is difficult to calculate K–entropy with (1). For practical purpose Grassberger et al. (1983) presented an approximate method. They rewrote (1) to q–order Renyi entropy, i.e.,

$$K_q = -\lim_{k\to 0}\lim_{\varepsilon\to 0}\lim_{d\to\infty}\frac{1}{dk}\frac{1}{q-1}\ln\sum_{i_1\cdots i_d} P^q(i_1,\cdots,i_d) \tag{2}$$

Obviously $K = \lim_{q\to 1} + K_q$. It also can be seen from (2) that as q increases, K_q decreases. Let $q=2$, then

$$K_2 = -\lim_{k\to 0}\lim_{\varepsilon\to 0}\lim_{d\to\infty}\frac{1}{dk}\ln\sum_{i_1\cdots i_d} P^2(i_1,\cdots,i_d) \tag{3}$$

where $K_2 < K$, the former is a lower bound of the latter. It has been proved by Grassberger et al. (1983) that K_2 retains major properties of K, so it can still act as a characterstic quantity to identify the feature of the system's motion in phase space. That means that with $K_2 > 0$, the system will take chaotic motion.

For further discussion, let

$$C_d(\varepsilon) = \lim_{N\to\infty}\frac{1}{N^2}\Big\{\text{with}\big(\sum_{i=1}^{d}|\vec{X}_{n+i} - \vec{X}_{m+i}|^2\big)^{\frac{1}{2}} < \varepsilon, \ \text{the number of points (n,m) } \textit{in pairs}\Big\} \tag{4}$$

Here $C_d(\varepsilon)$ is called the accumulative distance distribution function and can be given as

$$C_d(\varepsilon) = \lim_{N\to\infty}\frac{1}{N^2}\sum_{\substack{n,m=1\\ n\neq m}}^{N} Q(\varepsilon - r_{n,m}) \tag{5}$$

where $Q(g)$ is the Heaviside function with $g=(\varepsilon - y_{n,m})$; Q is zero for $g<0$ and one for $g \geqslant 0$.

The formula listed below has been obtained (Peng et al., 1989)

$$C_d(\varepsilon) \sim \varepsilon^D \tag{6}$$

where D denotes the associated dimension. Combining (3), (4) and (6), we have

$$C_d(\varepsilon) \sim \varepsilon^D e^{-K_2 dk} \tag{7}$$

If the phase space is increased from d dimenstion to $d+R$, we have

$$C_{d+R}(\varepsilon) \sim \varepsilon^D e^{-K_2(d+R)k} \tag{8}$$

From (7) and (8) we have

$$K_{2,d}(\varepsilon) = \frac{1}{Rk} \ln \frac{C_d(\varepsilon)}{C_{d+R}(\varepsilon)} \tag{9}$$

It is known that $C_d(\varepsilon)$ and $C_{d+R}(\varepsilon)$ are the probabilities of two adjoining points n and m being point pairs on the orbits of d and d+R dimension phase spaces, respectively, when $r_{n,m} \leqslant \varepsilon$. Thus, the formula (9) represents the mean divergence rate of a orbit section on which located the adjoining point pair n and m. As indicated in Yan et al. (1990), the Lyapunov exponent λ_i is a physical quantity for describing the variation of the geometrical features in the processes of contraction and expansion of the phase volume of a dissipative system in phase space, each positive λ_i denotes the feature of continuous expansion of a given system in a certain "direction", and the sum of all positive $\lambda_i s$ shows the overall expansion of the system. Interestingly, the value of K_2 given by (9) just reflects this feature, and both of them are correlated closely; it can be proved (Ruell, 1983) that

$$K_2 < K \leqslant \sum_{\lambda_i > 0} \lambda_i \tag{10}$$

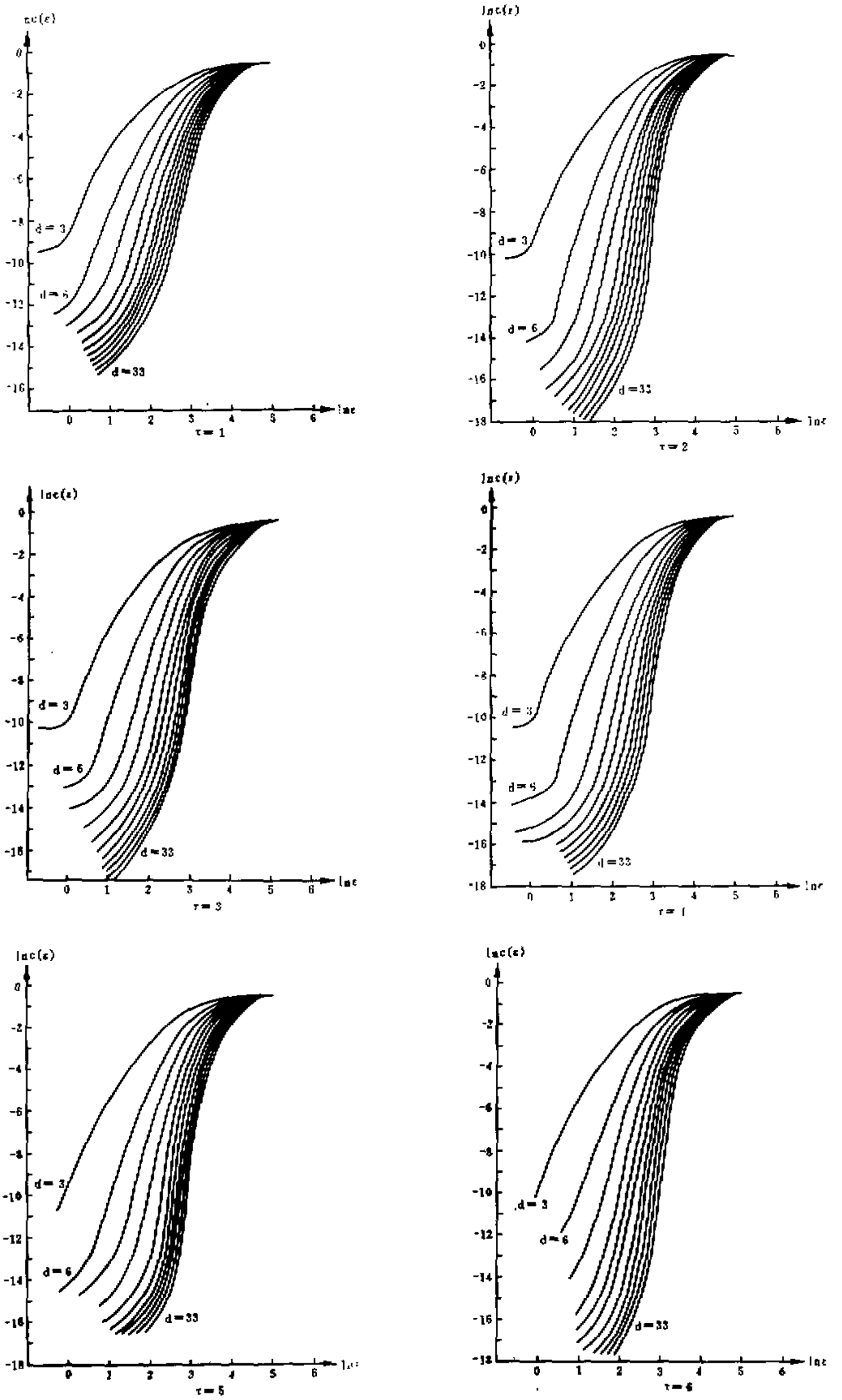
Owing to the expansion of the phase space orbit in certain directions, original adjacent states on attractor become no correlation leading to the long–term behaviour of the system being unpredictable. Thus we can see that either the K–entropy or the sum of the lyapunov exponents $\sum_{\lambda_i > 0} \lambda_i$ has ability to characterize the gross expansion of the system, therefore, to indicate the unpredictability the system has in its development process. On the other hand, the reciprocal of the sum can be used to characterize the time scale of the predictability, i.e., $T = \frac{1}{\sum_{\lambda_i > 0} \lambda_i} \sim 1/K$.

III. EXTRACTION OF K_2 FROM ONE–DIMENSIONAL TIME SERIES

The data of 1970–85 day to day averaged pressure of Shanghai, coming from the compilation of data of National Meteorological Bureau, were used to calculate the time series, the length of it is $N = 5000$.

The series are extended into d–dimensional phase space with the lag time $\tau = 1,2,3,\cdots\cdots,8$. The values of d are 3, 6, 9, 12, $\cdots\cdots$, 33. According to (5), each ε can have a $C_d(\varepsilon)$, the results are summarized in Fig.1a–h.

It can be seen from Fig.1 that in each of the curves with $\tau = 1,2,\cdots\cdots,8$ exists a saturation slope, and d_∞, being called *saturation imbedding dimension*, is the value of d corresponding to the saturation slope. The slope of the straght–line part of the curve was calculated, and the correlation dimension D of the system is


lnc(ε)
d=3
d=6
d=33
lnε
τ=1
τ=2
τ=3
τ=5
τ=6

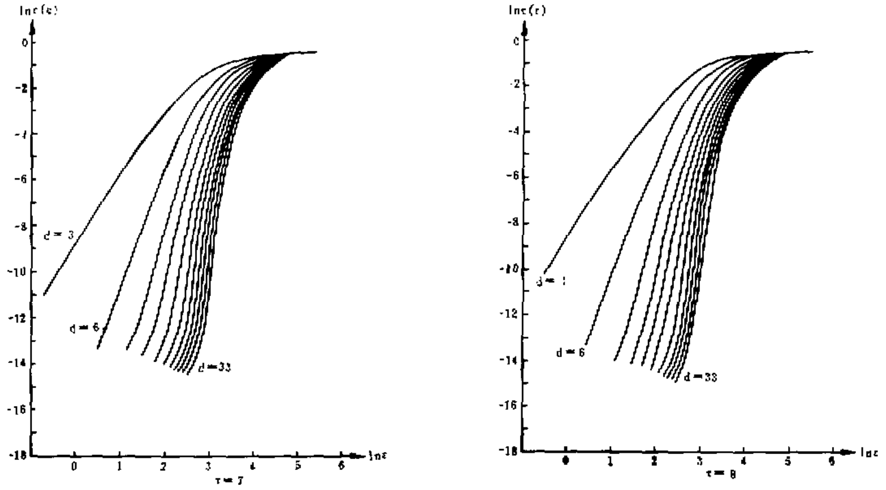


Fig.1. $C_d(\varepsilon)$ for different d–dimensional phase spaces.

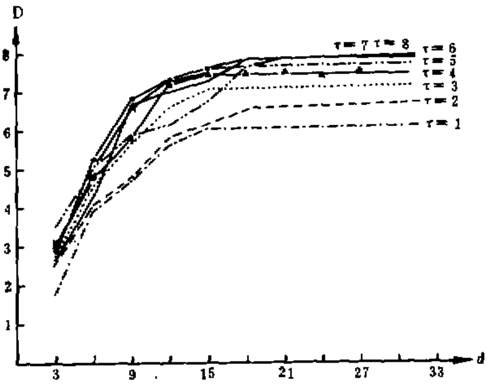


Fig.2. Relationship between D and d.

$$D = \frac{\ln C(\varepsilon_2) - \ln C(\varepsilon_1)}{\ln \varepsilon_2 - \ln \varepsilon_1} \tag{11}$$

The results are shown in Fig.2, and D and d_∞ obtained are listed in Table 1.

Table 1. The correlation dimension D and the saturation imbedding dimension d for different time lag τ

τ	1	2	3	4	5	6	7	8
D	6.1	6.7	7.1	7.5	7.7	7.8	7.9	7.9
d_∞	15	15	18	15	18	18	18	18

From Fig.1 and (9) with $R = 3$ and $k = 1$, corresponding $k_{2,d}(\varepsilon)$ for different ε can be obtained, and

$$\lim_{\substack{d\to\infty \\ \varepsilon\to 0}} K_{2,d}(\varepsilon) \sim K_2 \tag{12}$$

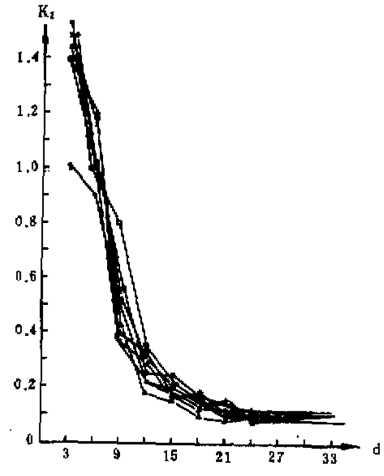


Fig.3. Relation between K_2 and d.

For small ε, the relationship beweeen K_2 and d is illustrated in Fig.3.

It is apparent in Fig.3 that K_2 tends to be a stable value with d's increasing, the stable value of K_2 and the time scale of predictability, $T=1/K_2$, for different τ are given in Table 2.

Table 2. Relations of K_2 and T to τ

τ	1	2	3	4	5	6	7	8	mean
K_2	0.1002	0.0831	0.1000	0.1166	0.1003	0.1001	0.1000	0.0934	0.0992
$T=1/K_2$	9.98	12.03	10.00	8.57	9.97	9.99	10.00	10.70	10.07

IV. DISCUSSIONS

Based on the data of Shanghai day to day pressure, the above calculations show the correlation dimension $D=7.7\sim7.9$ and the saturation imbedding dimension $d_\infty=18$. It is worth noticing that the correlation dimension $d_s=3,4$, which is for the evolution of short–term climate system, given in the previous paper (Peng et al., 1989) is obtained by using the data of monthly mean temperature of Shanghai. It is smaller comparing to $D=7.7\sim7.9$. Obviously it is due to that the background of monthly element describing the short–range climate change is different from the one of weather of the day to day averaged element. It is more complicated to describe daily mean variation of weather system than to describe monthly mean change of it. Thus, more variables are required in the first case.

It is evident from Table 2 that the mean value of K_2 is 0.0992 (a limted positive) which indicates that the background of the weather attractor is of chaotic motion in monsoon area. The predictability time scale obtained from K_2 is $T=1/K_2=10.07$ days. It is in

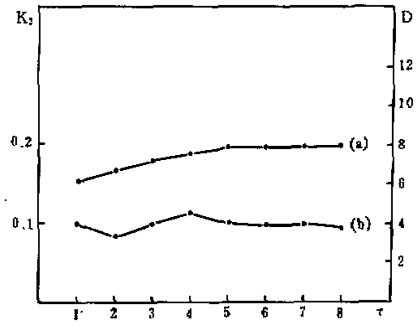


Fig.4. D and K_2 are convergent with respect to τ, (a) D to τ (b) K_2 to τ.

agreement with that obtained from dynamic– statistical method in early days, and is currently accepted as the predictable time scale of short–term weather prediction model. This consistency is meaningful, because the results in this paper are obtained based on the chaotic motion of the system. This kind of method is very simple, and can be extended to investigate observational data with various time scales and without meeting difficulties the dynamical statistical method has.

In this study the time lag τ is assigned to be 1, 2,…… and 8 in order to investigate the effects of different lags on the calcalution results. The dynamical characteristic quantities can be reliable values only if the components of the coordinate system are independent of each other in the phase space reconstracted with one dimensional time series. The results show that as τ grows, the correlation dimensions, which are convergent with respect to τ_1 increase as shown in Fig.4a.

For the data of day to day mean pressure used in this study, D tends to be 7.7~7.9 and the corresponding saturation imbedding dimension d_∞ is arourd 18 when τ is about 5.

It can be seen from the calculation recults of K_2 that the growth of τ causes slight fluctuation of K_2, but when $\tau > 5$, K_2 tends to be a stable value around 0.1 as shown in Fig.4b. This further demonstrates that in the reconstructed phase space with $\tau = 5$ the coordinate components are independent mutually and the characteristic quantity tends to be a stable value.

REFERENCES

Eckmann, J. P. (1985), Ergodic theory of chaos and strange attractors, *Rev. Mod. Phys.*, **57**: 617–655.

Grassberger, P. et al. (1983), Estimate of the Kolmogorov entropy from a chaotic signal, *Phys. Rev. A*, **28**: 2591–2593.

Peng Yongqing, Yan Shaojin and Wang Jianzhong (1989), Continuation of 1D climatic time series and determination of the number of dimension of the chaotic attracor in the phase space. *Trop. Met.*, **5**: 97–104.

Ruelle, D. (1983), Five turbulent problems. Physica 7D, 40–44.

State Meteorological Administration, PRC: Monthly bulletin of surface meteorological records. 1970–1985.

Wolf, A. and J. B. Swift et al. (1985), Determining Lyapunov exponents from a time series. Physica 16D 285–317.

Yan Shaojin, Peng Yongqing and Wang Jianzhong (1990), Evolutional feature of a dynamic system displayed by lyapunov exponents of 1D climatic time series, *Trop. Met.*, **6**: 105–112.