PHYSICAL REVIEW A VOLUME 25, NUMBER 6 JUNE 1982

# Impracticality of a box-counting algorithm for calculating the dimensionality of strange attractors

H. S. Greenside
*Bell Laboratories, Murray Hill, New Jersey 07974*

A. Wolf and J. Swift
*Department of Physics, University of Texas at Austin, Austin, Texas 78712*

T. Pignataro
*Joseph Henry Laboratories, Princeton University, Princeton, New Jersey 08544*

(Received 7 December 1981)

An algorithm proposed by Takens, which can determine the capacity (generalized dimensionality) of a dynamical system from the time series of a single observable, is tested numerically for several intrinsically stochastic models. The algorithm is found to converge too slowly (if at all) to be useful for the analysis of experimental data.

In a recent series of papers[1,2] Takens suggested an algorithm for computing the capacity of a dynamical system from the measurement of a single observable. The capacity is a particularly useful number to compute for experimental data taken near the onset of turbulence where the issue is to determine the number of degrees of freedom involved. If the capacity is a finite number, then the time dependence arises from a finite-dimensional deterministic mathematical model. If the capacity is nonintegral then the data are inconsistent with Landau's deterministic theory of turbulence[3] in which phase-space motion occurs on a high-dimensional torus. A theory based on a strange attractor[4] is then likely to be correct. On the other hand, an infinite capacity implies that an infinite number of degrees of freedom are needed to describe the dynamical system which would rule out both Landau's theory and the hypothesis of strange attractors.

We shall see below that there are severe computational difficulties in calculating the capacity of any set whose capacity is greater than two and that box counting by Takens's algorithm converges too slowly to be useful for the analysis of experimental data. The impracticality of using box-counting algorithms has been previously suggested in the literature,[5,6] but our calculation is the first to demonstrate quantitatively the limits of this approach. Takens's idea of reconstructing an attractor from a time series has also been independently proposed and studied,[6] but the emphasis was on obtaining the integral part of the capacity by fitting linear subspaces of appropriate dimension to sets of points on the reconstructed attractor. In the following, we describe Takens's algorithm and why the method converges so slowly.

The capacity of an attracting set $M$ which we assume is contained in an invariant manifold of some dynamical system, is easily calculated in principle if the coordinates of each point of $M$ in phase space are known. If $N_\epsilon$ is the minimum number of boxes of side $\epsilon$ in phase space needed to cover the set $M$, then, by definition, the capacity $C$ is given by the following limit[7]:

$$C = \lim_{\epsilon \to 0} \ln N_\epsilon / \ln(1/\epsilon) \quad . \tag{1}$$

Intuitively, if a set has volume $V$, the number of boxes of side $\epsilon$ needed to cover the set is roughly $N_\epsilon \cong V\epsilon^{-C}$ for sufficiently small $\epsilon$. This relation may be rewritten as

$$\ln N_\epsilon = C \ln(1/\epsilon) + \ln V \quad , \tag{2}$$

which provides a more practical method of computing $C$ than Eq. (1), since the latter has a slowly vanishing correction to the capacity $V/\ln(1/\epsilon)$.[5] By plotting $\log_{10} N_\epsilon$ against $\log_{10}(1/\epsilon)$ for decreasing values of $\epsilon$, Eq. (2) gives the capacity as the asymptotic slope (see Fig. 1).

The definition, Eq. (1), cannot be applied to experimental data which, typically, provide information about one phase-space coordinate only. Several ways of getting around this problem have been suggested[1,2,6] and, with the recent work of Takens, they have been placed on a rigorous foundation. Takens showed that for nearly all smooth dynamical systems it was possible to construct from a single-time series (observable) a new manifold whose essential properties (in particular, the capacity) are the same as those of the original manifold $M$. If the time series is represented by the infinite sequence of real numbers $\{a_i\}_{i=1}^{\infty}$ then the infinite set of $D = n+1$ dimensional vectors

$$S_D = \{\langle a_i, \ldots, a_{i+n} \rangle\}_{i=1}^{\infty}, \quad n \geqslant 0 \tag{3}$$

will give an embedding (a one-to-one differentiable map whose inverse is also differentiable[8]) of the original manifold for almost every choice of the observable, provided that the integer $n$ is not less than *twice* the capacity of the original manifold.[1] This requirement is actually much stronger than what is needed to calculate the capacity. If we relax this condition and only require that $n$ not be less than the capacity of $M$, the map, Eq. (3), will not necessarily be an embedding (e.g., one-to-one) but will generally suffice to determine the capacity.[9]

Takens's algorithm for computing the capacity is then the direct application of the box-counting method described by Eqs. (1) and (2) to the set of points given by Eq (3). For points $a_i$ sampled at equal-time intervals, the method is equivalent to that of Ref. 6 where $n+1$ delay coordinates are used to create a phase-space representation of the data and $n+1$ dimensional boxes of side $\epsilon$ are used to cover the representation. For a finite-time series of length $N$, $\{a_i\}_{i=1}^N$, $S_D$ will be a finite approximation of $M$, provided transients have been allowed to decay; one attempts to calculate the capacity of the $(N-n)$ vectors in $S_D$ for several increasing values of $N$ and for several choices of the integer $n$. For $N$ and $n$ sufficiently large ($n$ must be at least the capacity of $M$), the capacity of $S_D$ will be independent of both parameters and will give an approximation, whose accuracy depends on $\epsilon$, to the true capacity. The magnitude of the errors in the data (arising from noise, for example) must be smaller than the smallest $\epsilon$ used in taking the limit, Eq. (1), in order to obtain a meaningful estimate.

We emphasize that the most important part of Takens's analysis is that the set, Eq. (3), will almost always be an embedding which can then be used to obtain information about the manifold $M$. Given the set $S_D$, there are more or less straightforward ways to calculate the capacity.[1,5] If the set $S_D$ has a large capacity (say, greater than 3), then the numerical calculation is difficult no matter which technique is used.

We have tested Takens's algorithm on several sets and dynamical systems for which the capacity is known by other methods.[5] Our results are summarized in Table I and in Figs. 1 and 2. For low-dimensional sets ($C \leqslant 2$) the method works and a reasonable number of points suffice to determine the capacity to several significant figures. The dependence of the capacity of $S_D$ [Eq. (3)] on $n$ is shown in Fig. 1 for the Hénon map.[10] We see that for $n \geqslant 2$, the capacity is independent of $n$ and is about $1.26 \pm 0.01$, in excellent agreement with Ref. 5.

We have also applied the algorithm to two hydrodynamical models, a three-variable model by Lorenz[11] and its extension to 14 variables by Curry.[12] In a certain parameter range and for certain initial conditions, these models have intrinsically stochastic solutions whose corresponding strange attractors must have capacities greater than two. For sufficiently large $\epsilon$, Takens's method converged, but did not yield meaningful estimates of the capacity. For smaller $\epsilon$, the algorithm failed to converge for both models even when about a million points were used. The lack of convergence for the Curry model is

TABLE I. Comparison of Takens's and other methods to obtain the capacity.

| System | Capacity by Takens's method | Max number of points in series | Capacity by other means |
|---|---|---|---|
| $\frac{2}{3}$ Cantor set[a] | $0.63 \pm 0.02$ | 10 000 | 0.631 |
| Quadratic return map[b] | $0.98 \pm 0.02$ | 50 000 | 1.000 |
| Hénon map[c] | $1.26 \pm 0.01$ | 200 000 | $1.261 \pm 0.003$ |
| Lorenz model[d] | No convergence | 200 000 | $2.06 \pm 0.01$ |
| Curry model[d] | No convergence | 800 000 | $2.4 \pm 0.1$ |

[a]Defined recursively by $a_1 = 0$; $a_{2^k+i} = a_i + 2/3^{k+1}$, $1 \leqslant i \leqslant 2^k$ for $k \geqslant 0$. Reference 7 demonstrates that the capacity is exactly $\ln 2/\ln 3 \cong 0.6309$.
[b]The series is defined recursively by $a_1 = 0.2$, $a_{n+1} = 4a_n(1-a_n)$ for $n \geqslant 1$. The capacity can be analytically proven to be exactly one for nearly all starting conditions $0 < a_1 < 1$. See Ref. 14.
[c]See Ref. 10 for the definition and parameter values and Ref. 5 for the value given in the last column.
[d]See Ref. 11 for the Lorenz model and Ref. 12 for the Curry model. The numbers in the last column are not the capacity but the Lyapunov dimension (Ref. 13), which is a lower bound to the capacity and in a previous study (Ref. 5) found to be equal to the capacity. For a detailed comparison of these and other definitions of dimension, see D. Farmer, Ph.D. thesis (University of California at Santa Cruz, 1981) (unpublished).

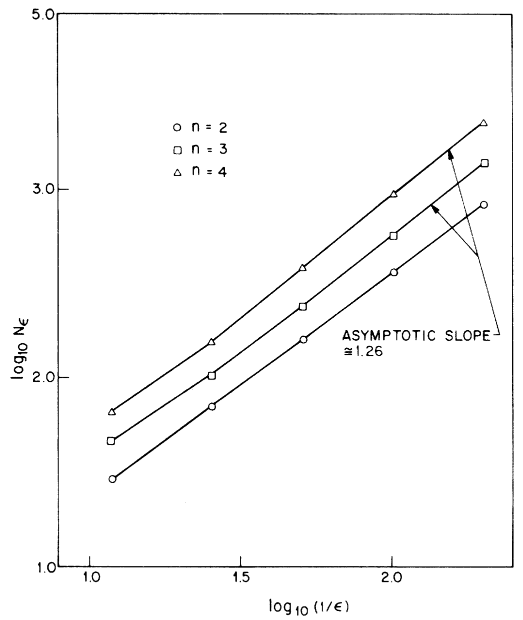


FIG. 1. Example of how Eq. (2) of the text can be used to calculate the capacity of a set, in this case the Hénon attractor (Ref. 10). The asymptotic slopes for $n = 2$, 3, and 4 [see Eq. (3) of the text] are 1.18, 1.26, and 1.26, respectively, to three significant digits. A time series containing 200 000 points was used for these results. The values of $\epsilon$ have been normalized to the range of the time series, the maximum value (1.27) minus the minimum value (−1.28).

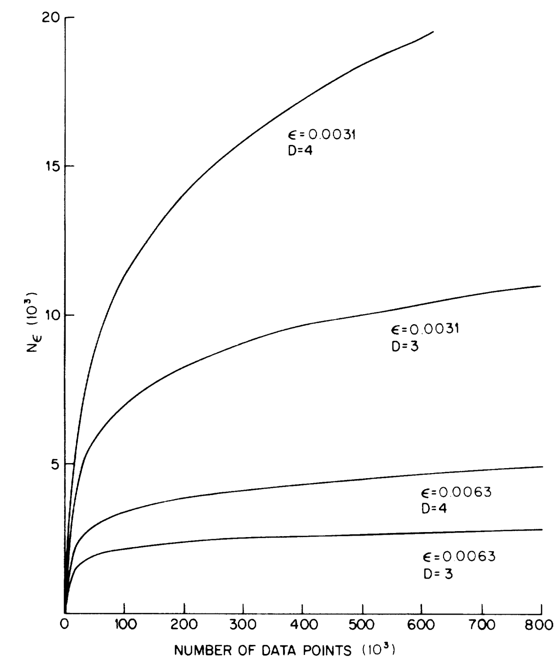


FIG. 2. Number of $D$-dimensional boxes of side $\epsilon$, $N_\epsilon$, that contain at least one element of the set $S_D$ [see Eq. (3) of the text] vs the number of points in the time series. The number $N_\epsilon$ fails to reach a constant value even for 800 000 points. The results are for the $\psi_{11}$ variable (which varies between −80 and +80) of the 14-variable Curry model (Ref. 12) in a parameter range which gives an intrinsically chaotic solution. The values of $\epsilon$ have been normalized to the range of the time series, the maximum value (80) minus the minimum value (−80).

shown in Fig. 2. For fixed $\epsilon$, $N_\epsilon$ does not stop growing as a longer and longer time series is used.

There are several reasons why enormously long-time series may be needed to calculate the capacity of a particular dynamical system by box-counting methods.[13] As has been mentioned several times in the literature,[5,6] the most fundamental is the exponential dependence of $N_\epsilon$ on the capacity [Eq. (1)] which poses a seemingly insurmountable difficulty for large capacity manifolds. A second reason is that the set $S_D$ often fills out the attractor in a highly nonuniform way. Many points are needed in $S_D$ to obtain those few points that provide a covering, especially for the parts of the attractor that are rarely visited. A simple example is a one-dimensional return map which ergodically fills out the unit interval, but with a vanishing probability density near 0 and 1.[14] A third reason, which applies to dissipative systems, is that the dynamical system may rapidly contract volumes in phase space, making it difficult to obtain the nonintegral part of the capacity which arises from the fractal structure.[7] The size of $\epsilon$ needed to resolve the fractal structure is then much smaller than a length scale which can cover the attractor in a reasonable number of points. The Lorenz and Curry attractors have high dimensionality, certain parts are visited rarely and nonuniformly, and they rapidly contract phase space so that they provide an extreme test of Takens's algorithm. On the other hand, the low dimensionality and slow rate of phase-space contraction for the $\frac{2}{3}$ Cantor set and the Hénon map (see Table I) facilitated rapid convergence to a nonintegral capacity.

A general statement about the number of points required to obtain convergence for a given $\epsilon$ cannot be made, considering the diverse possibilities for attractors. For the same reason, a general statement about how noise will affect the calculation to some given precision cannot be made. Our results for the Lorenz and Curry models suggest that for higher-dimensional manifolds an utterly impractical number of data points would be needed just to obtain the capacity to one significant figure.[15] The slow convergence of $N_\epsilon$ with the length of the time series suggests that the calculation of other useful numbers, such as the topological entropy,[1] or the calculation of the capacity by a probabilistic approach,[16] will also not be possible.

In summary, we have tested Takens's algorithm on several mathematical models. While we obtained meaningful results for sets contained in a two-dimensional phase space, the algorithm did not converge for higher-dimensional models even when nearly a million points in the time series were used. The essential reason for the slow convergence was not the manner in which the manifold was reconstructed [Eq. (3)] but the inherent difficulty of using box counting on a high-dimensional set. There might be experimental systems[17] of sufficiently low dimensionality and slow contraction that a box-counting algorithm could successfully be applied, especially if a Poincaré section were used to reduce the dimensionality by one (see Ref. 5). However, alternative ways of quantitatively comparing experimental data with deterministic models are needed.

## ACKNOWLEDGMENTS

The authors wish to thank N. Packard, D. Rand, D. Russell, F. Takens, and J. Yorke for several useful discussions and comments. We especially thank H. Swinney for providing encouragement and strong motivation for part of this project. Part of this research was supported by the Robert A. Welch Foundation under Grant No. F-767 and by the NSF under Grants No. DMR-76-11426 and CME-79-09585.

---

[1]F. Takens, in Proceedings of the Warwick Symposium, 1981, edited by D. Rand and L. S. Young (unpublished).

[2]F. Takens (unpublished).

[3]L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon, New York, 1959), Sec. 27.

[4]Although there is no universally accepted definition, we will define a strange attractor to be a bounded attracting set in the phase space of some deterministic dynamical system such that all nearby paths on the attractor diverge (which rules out limit points and limit cycles). For a good discussion, see R. Shaw, Z. Naturforsch. 36A, 80 (1981), and D. Ruelle, The Mathematical Intelligencer 3, 126 (1980).

[5]D. A. Russell, J. D. Hanson, and E. Ott, Phys. Rev. Lett. 45, 1175 (1980).

[6]N. H. Packard, J. P. Crutchfield, J. D. Farmer, and R. S. Shaw, Phys. Rev. Lett. 45, 712 (1980); H. Froehling, J. P. Crutchfield, D. Farmer, N. H. Packard, and R. Shaw, Physica (Utrecht) 3D, 605 (1981).

[7]B. B. Mandelbrot, *Fractals: Form, Chance, and Dimension* (Freeman, San Francisco, 1977).

[8]Y. Choquet-Bruhat, C. de Witt-Morette, M. Dillard-Bleick, *Analysis, Manifolds, and Physics* (North-Holland, Amsterdam, 1977).

[9]A simple example (suggested to us by F. Takens) is the mapping of the unit circle into the plane as a figure "eight" ($\infty$). This map is not an embedding nor are small perturbations of this map (the map is not one-to-one). A map of the unit circle into $2n+1=3$ dimensions will generally be an embedding since small perturbations "open up" the figure eight.

[10]M. Hénon, Commun. Math. Phys. 50, 69 (1976). We used the parameter values $a=1.4$, $b=0.3$ with initial conditions $x_0=0.2$, $y_0=0.2$. The first $10^5$ iterations were discarded to allow transients to disappear and the value of the $x$ variable was used for the time series.

[11]E. Lorenz, J. Atmos. Sci. 20, 130 (1963). We used the parameters $r=28$, $\sigma=10$, $b=\frac{8}{3}$ with initial conditions $x_0=y_0=z_0=1$. A sampling rate of 0.003, 0.012, and 0.03 time units was used on the $x$ variable to form the time series, the latter value corresponding to about 30 points per cycle. The first $10^4$ points were discarded to avoid transients.

[12]J. Curry, Commun. Math. Phys. 60, 193 (1978). We used the parameter values $a=1/\sqrt{2}$, $\sigma=10$, and $r=45.92$ with a sampling rate of 0.012 time units; the first $10^4$ points were discarded to avoid transients. The equations listed in this paper have been printed incorrectly. The correct equations are available from Dr. Curry and from the authors.

[13]One can also calculate the dimension of a dynamical system by a method using the Lyapunov exponents [see Ref. 5; P. Frederickson, J. Kaplan, E. Yorke, and J. Yorke (unpublished); R. Snapp and A. Wolf (unpublished)]. Although this method converges much more rapidly than box-counting methods, it requires knowledge of the Jacobian of the dynamical system at each point in time and is therefore difficult to apply to experimental data of limited precision.

[14]S. Grossman and S. Thomae, Z. Naturforsch. 32A, 1353 (1977).

[15]Besides the huge number of data points needed, one cannot ignore the large and expensive amount of computer time needed to effect Takens's algorithm, even on a powerful computer. Straightforward FORTRAN programs on a CDC Cyber 170 computer required several hours of processing time to generate and analyze the 800 000 point time series studied for the Curry model. For box-counting algorithms, the computational effort grows exponentially with the dimension of the system. (For the Lyapunov-exponent algorithms, the effort grows linearly with the dimension.)

[16]After this paper was nearly finished, F. Takens informed us of a second probabilistic algorithm that could be used to calculate the capacity. Preliminary calculations using his new approach are not encouraging; the algorithm generates a sequence of estimates for the capacity which do not converge. For example, using a 40 000 point time series for the Henon map ($a=1.4$, $b=0.3$, $C=1.263$), the values of the capacity were 1.2761, 1.2634, 1.2522, and 1.2500 for $\epsilon=0.05$, 0.025, 0.0125, and 0.00625, respectively. The new algorithm is not only subject to the same slow convergence as the first but has the disadvantage of requiring more computer memory than the one discussed in the text.

[17]J. C. Roux, A. Rossi, S. Bachelart, and C. Vidal, Phys. Lett. 77A, 391 (1980); J. S. Turner, J. C. Roux, W. D. McCormick, and H. L. Swinney, *ibid.* 85A, 9 (1981).