An introduction to chaotic signal processing

Article *in* GEC Journal of Research · January 1992

CITATIONS
8

READS
845

2 authors, including:

Michael J. Kearney
University of Surrey
138 PUBLICATIONS **1,804** CITATIONS

SEE PROFILE

AN INTRODUCTION TO CHAOTIC SIGNAL PROCESSING

M. J. KEARNEY and J. STARK

We present an overview of the basic concepts of chaotic time series analysis. We show how to model the underlying dynamics which gives rise to such time series from the measurement of a single observable. This can be used to devise novel schemes for prediction, noise reduction and feature extraction in temporal signal processing. Time series derived from simple non-linear dynamical systems are used to illustrate the underlying concepts.

Keywords: signal processing; non-linear modelling; chaos.

1. INTRODUCTION

Conventional signal processing or time series analysis has been limited for many years by the underlying assumption of linearity. In the real world, of course, this assumption is often far from reasonable, and much effort has recently been expended in trying to extend our understanding of non-linear systems. Central to this has been the realization that even very simple non-linearities can lead to extremely complex behaviour. The most dramatic manifestation of this is deterministic chaos, where the observed behaviour is apparently 'random' (exhibiting a continuous power spectrum reminiscent of coloured noise).

Chaos occurs naturally in many different non-linear mechanical, electrical, and electronic systems. Until very recently, however, it has tended to be modelled as a form of 'noise' because of the popular misconception that chaos is 'totally unpredictable'. In this paper we shall review a number of recent developments which show that this is not the case. In fact, we shall see that it is sometimes possible to predict chaotic sequences with great certainty over short timescales. Although we have chosen chaotic systems to illustrate a specific point, namely that non-linear (as opposed to linear) modelling conveys certain advantages, the methods and algorithms outlined can be used as a practical aid for the modelling of almost any non-linear process. It is not our intention to be too mathematical, or too detailed, and those who are interested in learning more are referred to the bibliography presented at the end.

2. IN WHAT SENSE ARE CHAOTIC SYSTEMS 'RANDOM'?

The best way to understand how a deterministic system can give rise to an element of 'unpredictability' (chaos) is through the notion of sensitivity to initial conditions. To explain what this means we consider a very simple (one-dimensional) chaotic system, namely the discrete mapping $x_{n+1} = f(x_n)$, with $0 \leq x_n \leq 1$, where f is given by

$$f(x) = \begin{cases} 2x & \text{if } 0 \leq x \leq 0.5 \ , \\ 2x-1 & \text{if } 0.5 < x \leq 1 \ . \end{cases} \qquad (1)$$

This mapping is non-linear by virtue of the discontinuity at $x = 0.5$. However, if we regard x_n as an angular variable and set $\theta_n = 2\pi x_n$, then the dynamics obtained from eqn. (1) is equivalent to that given by $\theta_{n+1} = 2\theta_n$ and the discontinuity disappears. Note that time evolves here in discrete steps, indexed by the integer $n = 0, 1, 2, \ldots$. This is quite natural within the context of time series analysis. An initial value x_0 (with $0 \leq x_0 \leq 1$) is chosen and substituted into eqn. (1) to generate the next iterate x_1, which is used to generate x_2, which generates x_3 and so on. The initial point x_0 thus determines the whole time series $\{x_n : n = 0, 1, 2, \ldots\} = \{x_0, x_1, x_2, \ldots\}$. To understand the behaviour of this dynamical system we write the initial value x_0 in its binary representation,

$$x_0 = \sum_{i=1}^{\infty} a_i(\tfrac{1}{2})^i \equiv [a_1, a_2, a_3, \ldots] \ , \qquad (2)$$

where the a_i are all either 0 or 1. For the sake of argument let us choose $x_0 \cong 0.5903320\ldots$, or in its binary notation (given to 12-bit accuracy),

$$x_0 = [1,0,0,1,0,1,1,1,0,0,1,0,*,*,\ldots].$$

The *s denote higher order bits that we do not know. One can easily show from eqn. (1) and eqn. (2) that the subsequent terms in the sequence are given by

$$x_1 = [0,0,1,0,1,1,1,0,0,1,0,*,*,*,\ldots\ ],$$

$$x_2 = [0,1,0,1,1,1,0,0,1,0,*,*,*,*,\ldots\ ],$$

$$x_3 = [1,0,1,1,1,0,0,1,0,*,*,*,*,*,\ldots\ ],$$

$$x_4 = [0,1,1,1,0,0,1,0,*,*,*,*,*,*,\ldots\ ],$$

and so on. The pattern which emerges is that to go from x_n to x_{n+1} we delete the most significant bit in the binary expansion of x_n and shift all the subsequent bits one place to the left. Notice that the number of known bits is decreasing all the time. Thus 'information' about our starting point is being lost with each iteration. In

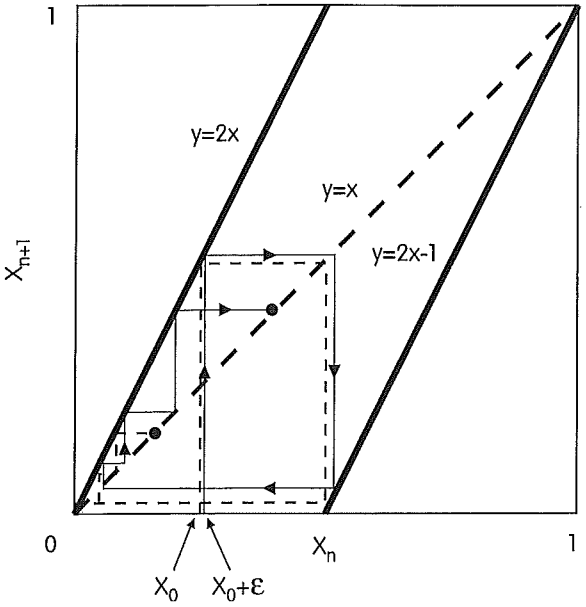


Fig. 1. Action of map in eqn. 1

our example, all the bits we know are deleted after 12 iterations, meaning that we have absolutely no idea where the iterate x_{12} lies in the interval [0,1]. Therefore, although we can predict ahead reasonably well for short timescales (a few iterations), on longer timescales accurate prediction becomes impossible. In the same spirit, the 'separation' between any two 12-bit numbers which, say, initially differ in only their last bit, will increase by a factor of 2 after each iteration.

More generally, consider two initial starting values x_0 and $x_0 + \epsilon$ which are close together (see fig. 1). After one iteration, $x_0 \rightarrow f(x_0) = x_1$ and $x_0 + \epsilon \rightarrow f(x_0 + \epsilon) = x_1 + 2\epsilon$. After two iterations $x_0 \rightarrow f(f(x_0)) = x_2$ and $x_0 + \epsilon \rightarrow f(f(x_0 + \epsilon)) = f(x_1 + 2\epsilon) = x_2 + 4\epsilon$. After n iterations the separation is $2^n\epsilon$ (assuming that initially ϵ was sufficiently small). We therefore see that small initial errors grow rapidly and points initially close together rapidly 'separate'. This effect is what is meant by 'sensitivity to initial conditions'. One of its most important consequences is that it is impossible to predict the long term behaviour of such a system despite the fact that it is completely deterministic.

Such sensitivity to initial conditions also has profound implications for the dynamics of the system. Consider the initial condition x_0 which gives rise to the sequence $\{x_n\}$. A typical x_0 will have an irregular sequence of 0s and 1s in its binary expansion and hence will lead to irregular (chaotic) behaviour. Exceptional values of x_0 will have regular periodic expansions and will give rise to periodic behaviour. Thus, for instance, $x_0 = [1,0,1,0,1,0,...]$ corresponds to the trajectory $x_0 = 2/3$, $x_1 = 1/3$, $x_2 = 2/3$, $x_3 = 1/3$, ... and so on. Whilst in principle this is an exact trajectory of eqn. (1), in practice, any slight deviation (caused by noise or rounding error in a computer) will rapidly grow and lead to aperiodic behaviour. It is therefore unlikely that such periodic behaviour would ever be observed in a real system and every sequence† $\{x_n\}$ will always appear to be 'random'. In fact, one can show that the statistical properties of $\{x_n\}$ are completely indistinguishable from those of purely random sequence (even when x_0 is known to arbitrary precision so that the sequence is, in theory, completely deterministic). Thus it will have a continuous power spectrum and if we simply note whether x_n is to the left or right of 1/2, the resulting left/right sequence is equivalent (as shown by all known statistical tests) to that obtained by tossing an unbiased coin.

Exponential divergence of nearby trajectories is a characteristic feature of all chaotic processes. It can be quantified for a general dynamical system through the notion of the Liapunov exponent λ. As above, consider two initial starting values x_0 and $x_0 + \epsilon$ which are close together. After one iteration, $x_0 \rightarrow f(x_0) = x_1$ and $x_0 + \epsilon \rightarrow f(x_0 + \epsilon)$, but now we do not necessarily have an explicit formula for $f(x_0 + \epsilon)$. However, if ϵ is sufficiently small, we can approximate this by $f(x_0 + \epsilon) \cong f(x_0) + f'(x_0)\epsilon = x_1 + f'(x_0)\epsilon$, where f' is the derivative of f. After two iterations $x_0 \rightarrow f(f(x_0)) = x_2$ and $x_0 + \epsilon \rightarrow f(f(x_0 + \epsilon)) \cong f(x_1 + f'(x_0)\epsilon) \cong x_2 + f'(x_1)f'(x_0)\epsilon$. After n iterations the separation will thus be approximately $f'(x_{n-1}) \ldots f'(x_1) f'(x_0)\epsilon$. If we are looking for exponential growth in the distance between the two trajectories, we want to find a λ such that after n iterations we have

$$|f^n(x_0 + \epsilon) - f^n(x_0)| \cong \epsilon \exp(n\lambda) , \qquad (3)$$

where f^n denotes f composed with itself n times. It follows from above that

$$\lambda = \lim_{n\rightarrow\infty} \frac{1}{n} \sum_{i=0}^{n-1} \ln |f'(x_i)| , \qquad (4)$$

assuming that the limit exists (which it does for almost all initial x_0). Observe that in the example given by eqn. (1) we have $f'(x) = 2$ for all x and hence $\lambda = \ln 2$.

A positive Liapunov exponent indicates exponential separation of nearby trajectories and hence is a sign of chaos. A negative exponent, on the other hand, shows that trajectories converge, and implies that the system is not chaotic.

3. EXTRACTING THE UNDERLYING DYNAMICS FROM A TIME SERIES

It should now be clear that the statement: 'chaos is totally unpredictable' is not strictly true. Chaotic sequences cannot be predicted over very long timescales (because we can only specify initial conditions to finite accuracy), but they can be predicted with some certainty over short timescales, provided that we know what the underlying dynamical system is. In the above discussion we only considered systems with one degree

† Note, however, that most computers will fill the least significant digits with zeros when they perform the multiplication by two. In that case, all trajectories will eventually end up at 0. Such behaviour is an artifact of the way that the computer performs arithmetic and is highly atypical of the dynamics of a real system.

of freedom. Real systems, however, are likely to have several degrees of freedom (possibly a very large number). Their state space will thus be multi-dimensional and the future behaviour of the system will be determined by a state vector rather than just a single variable. In many cases we have no way of directly determining this state vector and can merely take measurements of one (or possibly a few) observable quantities. For simplicity we shall restrict ourselves to just one observable. If this is measured at regular intervals it will give rise to a time series, which we shall continue to denote as $\{x_n\}$.

The question now arises as to whether or not we can use the information that $\{x_n\}$ has arisen from a deterministic system. In many cases the answer turns out to be 'yes' and, in fact, it often proves possible to recover an approximation of the underlying dynamical system. This can then be used to predict the (short-term) future behaviour of the time series. To illustrate how, we shall study a slightly more complicated (two-dimensional) system known as the Hénon map. This is given by

$$x_{n+1} = 1 - 1.4(x_n)^2 + y_n, \text{ and} \qquad (5a)$$

$$y_{n+1} = 0.3x_n . \qquad (5b)$$

Both the sequences $\{x_n\}$ and $\{y_n\}$ are chaotic and have continuous power spectra. A plot of 1000 values of $\{x_n\}$ is shown in fig. 2. Fig. 3(a) shows a plot of the position of successive pairs (x_n, y_n) in the x–y plane. The peculiar 'object' on which they lie is called a 'strange attractor'. Although it looks like a folded one-dimensional curve, it actually has a self-similar or 'fractal' structure. Thus, if we look at any part of it in detail, we see more and more structure at smaller and smaller scales (see fig. 3(b) which shows an expanded view of the box in fig. 3(a)). Such an object can be characterized by a non-integer 'dimension' which will be further discussed below.

If we knew the state (x_n, y_n) we could generate both x_{n+1} and y_{n+1} using eqn. (5). But suppose that we were only given the sequence $\{x_n\}$, and had no knowledge of $\{y_n\}$. Can we still hope to model the dynamics? By substituting eqn. (5b) into eqn. (5a), we can explicitly eliminate y_n to give

$$x_{n+1} = 1 - 1.4(x_n)^2 + 0.3x_{n-1}, \qquad (6)$$

and hence, in this case, the answer is clearly 'yes'. One way of thinking about this is that it gives some sort of second-order non-linear auto-regressive (AR) model for the sequence $\{x_n\}$. Eqn. (6) can be used to generate and hence predict the sequence $\{x_n\}$ from knowledge of x_0 and x_1. We can thus replace knowledge of the other state variable y_n by information about the previous behaviour of x_n (that is, in this case, the value of x_{n-1}).

The reason that this is possible in this example is the very special form of eqn. (5) and, in particular, the fact that y_{n+1} does not explicitly depend on y_n. In general, however, this will not be true and it would appear to be very unlikely that we can determine x_{n+1} from only a finite number of previous values of x_n without needing any information about y_n. Remarkably, this intuition turns out to be false. The Takens Embedding Theorem (named after its originator) states that, for all typical time series obtained from a finite degree of freedom

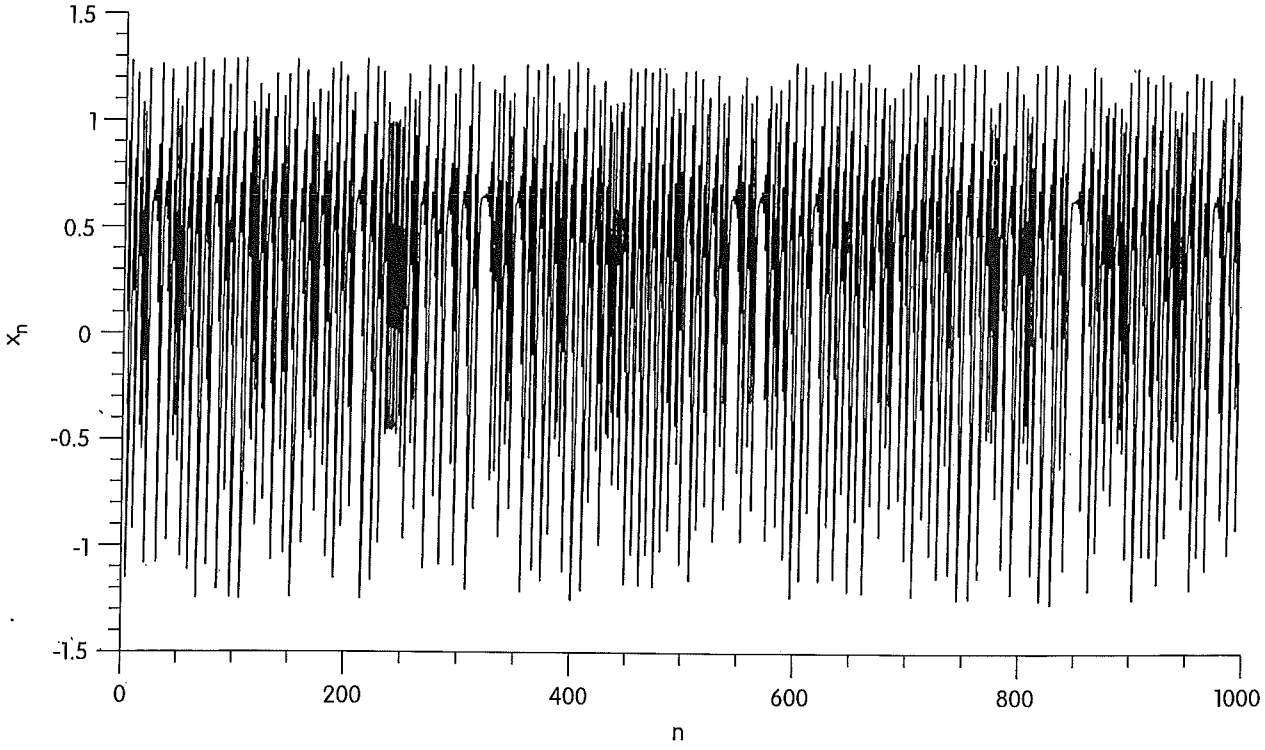
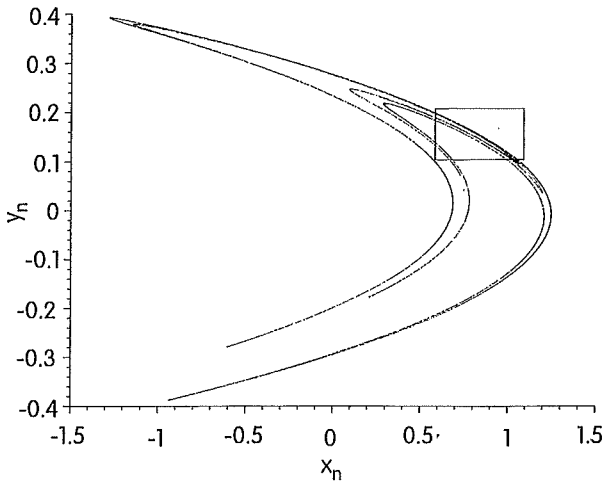
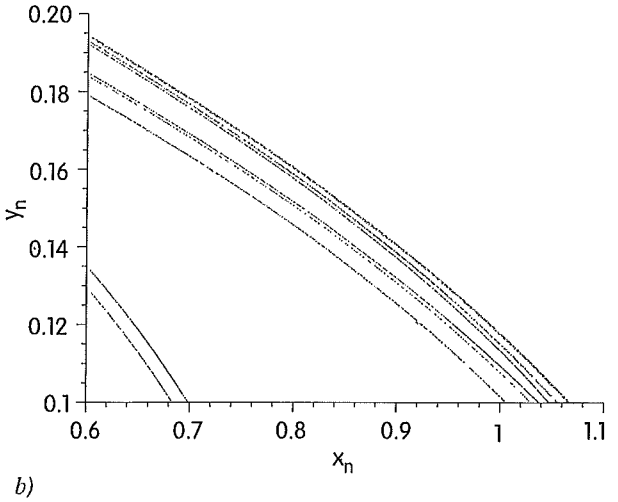


Fig. 2. Time series derived from the Hénon map in eqn. 5



a)



b)

Fig. 3. a) Attractor for the Hénon map in eqn. 5, and b) magnified view of box in fig. 3a)

dynamical system, there is some finite integer m and a function G such that

$$x_{n+1} = G(x_n, x_{n-1}, \ldots, x_{n-m+1}) . \qquad (7)$$

Furthermore, the dynamics generated by G is equivalent (up to a coordinate change) to the original dynamics which gave rise to $\{x_n\}$. In contrast to traditional AR models where G would be a linear combination of $x_n, x_{n-1}, \ldots, x_{n-m+1}$, the map G here will typically be non-linear. Thus in the above example $G(x_n, x_{n-1}) = 1 - 1.4(x_n)^2 + 0.3x_{n-1}$. The integer m is called the embedding dimension and can always be taken such that $m \leq 2d + 1$ where d is the dimension (number of state variables) of the underlying dynamical system. For the Hénon map (where $d = 2$) a value of $m = 2$ is sufficient, but for any typical two-dimensional map no more than 5 consecutive elements of a sequence are needed to determine the next element in that sequence.

This is an extremely powerful result. Unfortunately, Takens' theorem merely ensures the existence of some G and m, but does not tell us what the map G actually is; nor does it specify the value of m. As we shall see below (section 5), G can be readily estimated from the time series $\{x_n\}$ using standard non-linear approximation techniques. The value of m, on the other hand, has to be obtained essentially by trial and error. One possibility is to compute a trial G for each value m and then use the smallest m which gives a good fit to eqn. (7). A better method is to look at some simple invariants of a dynamical system which characterize its complexity and also give a good estimate of m as a by-product.

4. IDENTIFYING CHAOS IN A TIME SERIES

When given a complex-looking time series $\{x_n\}$, such as fig. 2, how do we tell whether it has come from a deterministic chaotic system or from a purely 'random' process? A variety of techniques for answering this question exist, all based upon the Takens embedding described above. Thus from the scalar series $\{x_n\}$ we form the vector series $\{\mathbf{v}_n\}$ where $\mathbf{v}_n = (x_n, x_{n-1} \ldots, x_{n-m+1})$. This process is called 'embedding' and will be familiar to some under the name of 'method of delays'. As mentioned above, the coordinate independent properties of $\{\mathbf{v}_n\}$ are the same as those of the original (unknown) dynamical system which gave rise to $\{x_n\}$. These include the Liapunov exponent λ described above, as well as a quantity called the 'correlation dimension', D_C. In principle it is possible to compute both λ and D_C from $\{\mathbf{v}_n\}$ (and hence from $\{x_n\}$), but in practice the calculation of D_C is much easier.

The correlation dimension seeks to measure the dimension of the set in m-dimensional space on which the points $\{\mathbf{v}_n\}$ lie. In some sense it is thus a measure of the complexity or number of variables required to describe this set. If all the $\mathbf{v}_n$ are identical (so that all the x_n are constant) D_C will be zero. If they lie on some curve then D_C will equal 1, and if they fill a plane it will equal 2. At the other extreme, if the $\mathbf{v}_n$ completely fill the m-dimensional space in which they lie we will have $D_C = m$. Intriguingly, D_C need not be an integer and is thus an example of a 'fractal dimension'.

We can estimate D_C from a finite sample $\{\mathbf{v}_1, \ldots, \mathbf{v}_N\}$ of our series of points as follows. First, form all the N^2 possible pairs $(\mathbf{v}_i, \mathbf{v}_j)$ of such points. Calculate the Euclidean distance $r_{ij} = |\mathbf{v}_i - \mathbf{v}_j|$ between each pair. Note that this can be defined in terms of the original $\{x_n\}$ by

$$r_{ij} = \left[\sum_{k=0}^{m-1} (x_{i-k} - x_{j-k})^2\right]^{1/2} . \qquad (8)$$

Now, for a given ϵ, let $N(\epsilon)$ be the number of pairs such that $r_{ij} \leq \epsilon$. Then $C(\epsilon) = N(\epsilon)/N^2$ is the proportion of pairs of points within a distance of ϵ of each other. A simple calculation shows that if all the points $\mathbf{v}_n$ lie randomly on some curve we have roughly $C(\epsilon) \sim \epsilon$ for large N and small ϵ. Similarly, if the $\mathbf{v}_n$

lie on a surface, we get $C(\epsilon) \sim \epsilon^2$. This suggests that $C(\epsilon)$ behaves exponentially for small ϵ with an exponent that is the dimension of the set on which the v_n lie. We thus define the correlation dimension by

$$D_C = \lim_{\epsilon \to 0} \{ \log C(\epsilon) / \log \epsilon \} . \quad (9)$$

The quantity $C(\epsilon)$ is called a correlation integral and several efficient numerical methods exist for calculating it, even for moderately large data sets (such as $N = 10^6$). To evaluate D_C we then plot $\log C(\epsilon)$ against $\log \epsilon$ and estimate the resulting slope.

So far we have assumed that we know the value of m required to embed our time series adequately. When this is not the case, we have to compute a correlation dimension $D_C(m)$ for for each choice of m. When m is too small, the set $\{v_n\}$ will completely fill the embedding space and we will get $D_C(m) \cong m$ (note that we always have $0 \leq D_C(m) \leq m$). But once m is sufficiently large, the computed value $D_C(m)$ will stabilize at the actual correlation dimension of the dynamical system. As an example, for the Hénon map above we get $D_C(1) \cong 1$ and $D_C(2) \cong D_C(3) \cong D_C(4) \cong 1.21$...which is the correct dimension of the Hénon attractor.

When D_C is not an integer, as in this case, we say that the system contains a 'strange attractor'. This is usually a sign of chaos (as indicated by a positive Liapunov exponent), although strange non-chaotic systems do exist (but are currently believed to be pathological).

It may happen that $D_C(m)$ continues to grow with m. This is an indication that the time series $\{x_n\}$ was generated by a truly random process, rather than a chaotic deterministic one. Indeed, for a random process such as white noise we have the rigorous result that $D_C = \infty$. By contrast, for a regular periodic time series we have $D_C = 0$ and, broadly speaking, as D_C increases we expect to see more and more complex behaviour.

The above procedure for computing D_C yields as a by-product the minimal embedding dimension required to represent adequately the dynamics of the system. It is thus the first step in evaluating the applicability of chaotic time series methods to a particular data set. In practice, lack of data and numerical precision limit calculations to about $m \leq 10$ (and hence $D_C \leq 10$). From the point of view of chaotic time series analysis, therefore, any system with D_C much larger than 10 is indistinguishable from a truly random one.

5. MODELLING THE DYNAMICS

The next step, after we have obtained a value for m using the above procedure (or by some other means), is to construct a model of the embedded dynamics. In other words, we want to construct an approximation to the function G such that $G(x_n, x_{n-1}, \ldots, x_{n-m+1}) = x_{n+1}$. This is a problem in non-linear function fitting and a number of different techniques can be used. We shall describe one, based upon radial basis functions, and illustrate its application to the Hénon map.

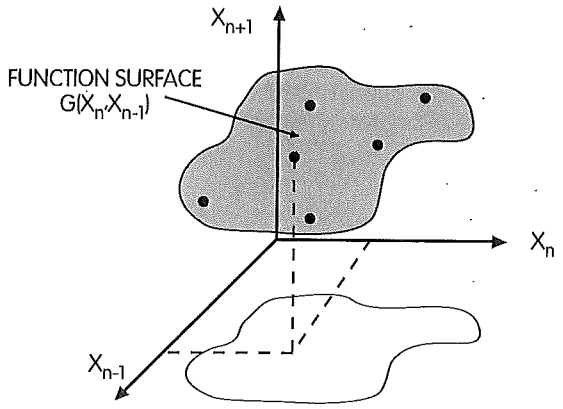


Fig. 4. The surface defined by the function G(X,Y)

For the Hénon map we have $m = 2$. We thus seek a function of two variables $G(X,Y)$ whose value at $X = x_n$, $Y = x_{n-1}$ is x_{n+1} for all n. In this particular example we happen to know that $G(X,Y) = 1 - 1.4X^2 + 0.3Y$, but let us suppose that we are unaware of this and are merely given a sample of the time series $\{x_n\}$. We can represent G as a surface $Z = G(X,Y)$ in three dimensional space as in fig. 4.

If we take our time series and plot the points $(X,Y,Z) = (x_n, x_{n-1}, x_{n+1})$ for all values of n, then they will all lie on this surface. Our search for an approximation to G is therefore equivalent to finding a smooth surface which passes through (or, at least, close to) all our sample points (x_n, x_{n-1}, x_{n+1}).

The radial basis approach approximates this surface as

$$G(X,Y) = \sum_{i=1}^{N-1} \lambda_i \phi\{[(X - x_i)^2 + (Y - x_{i-1})^2]^{1/2}\}, \quad (10)$$

where $\{x_0, \ldots, x_N\}$ is a given sample of the time series, the λ_i are unknown parameters and $\phi(r)$ is the so-called radial basis function. A convenient choice for this is $\phi(r) = (r^2 + c)^{1/2}$ for some constant $c > 0$, but many other choices are possible. The values of λ_i are determined by the requirement that $G(x_n, x_{n-1}) = x_{n+1}$ for $n = 1, \ldots, N-1$. This gives the linear matrix equation

$$z = \Phi\Lambda , \quad (11)$$

where z is the vector $z = (x_2, \ldots, x_N)$, Φ is the matrix whose $(i, j)^{\text{th}}$ term is $\Phi_{ij} = \phi(r_{ij})$ where r_{ij} is the Euclidean distance between the points (x_i, x_{i-1}) and (x_j, x_{j-1}) (computed as in eqn. (8) with $m = 2$) and Λ is the vector $\Lambda = (\lambda_1, \ldots, \lambda_{N-1})$. We can solve eqn. (11) for Λ by inverting Φ using any standard technique for the solution of linear equations.

To illustrate this procedure we used the first 72 points $\{x_0, \ldots, x_{71}\}$ of the Hénon time series in fig. 2 to construct an approximation (thus $N = 71$ in eqn. (10)). This was then evaluated in two ways. In fig. 5 we attempt to predict the time series forward from the 71st

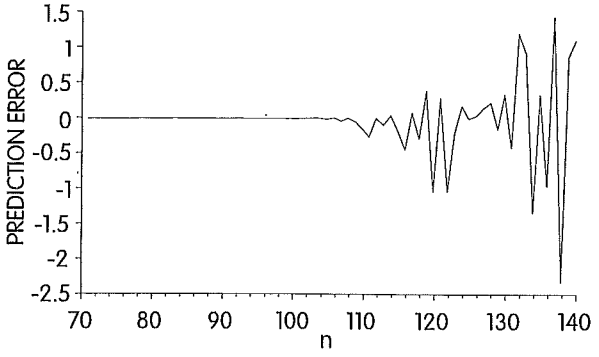


Fig. 5. Forecasting error in Hénon time series for repeated predictions forward from (x_{71}, x_{70})

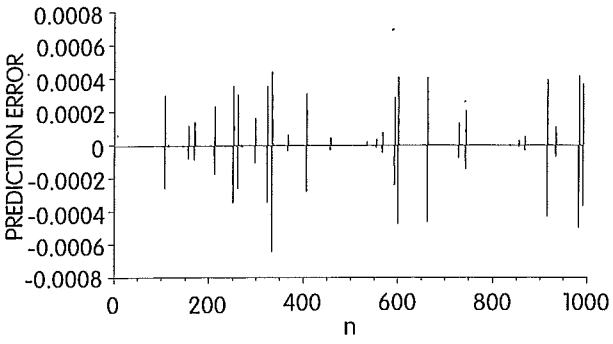


Fig. 6. One-step prediction error in Hénon time series

point. Thus the approximate G is applied to (x_{71}, x_{70}) to yield x_{72}, then to (x_{72}, x_{71}) (where x_{72} is the estimated value) to give x_{73} and so on. These predictions are then compared with the real time series. As we can see, the prediction is extremely accurate for about the first 35 time steps and then diverges drastically. This is in agreement with our remarks in section 2 that chaotic systems are predictable for short time scales but not for long timescales.

In the second test we simply used the actual time series data pairs (x_n, x_{n-1}) to predict one time step ahead and compare with the correct value x_{n+1} for the whole time series of fig. 2. Fig. 6 shows the result and, as we can see, we obtain excellent results throughout, with the relative error everywhere $\leq 10^{-3}$. More sophisticated implementations of the radial basis function approach can reduce this to $\leq 10^{-5}$.

One disadvantage of the above procedure is that the λ_i have to be calculated once and for all and this computation can be quite time consuming. As a result, the above algorithm is not particularly suited for real-time application. Recent work in the Systems Theory Group at the Hirst Research Centre, however, has led to the development of recursive versions of the above scheme that continuously update the λ_i using incoming values of the x_n, giving an algorithm much more appropriate to on-line forecasting.

Finally, note that the whole radial basis function framework can easily be extended to embedding dimensions m greater than 2. Thus, one forms the m-dimensional points $(x_i, x_{i-1}, \ldots, x_{i-m+1})$ (called 'centres') from the time series, and the argument of ϕ in eqn. (10) is simply the Euclidean distance between the point at which we want to evaluate G and the relevant 'centre'.

6. SIGNAL EXTRACTION AND NOISE REDUCTION

In most practical applications we are unlikely to be given a pure chaotic signal such as in fig. 2. Instead, we are asked to work with a mixture $u_n = x_n + s_n$ of a chaotic time series $\{x_n\}$ and some other signal $\{s_n\}$. The latter may represent noise, in which case we want to remove it from $\{u_n\}$, or it may be a signal that we wish to detect, in which case we want to extract it from $\{u_n\}$ and discard $\{x_n\}$. An example of the latter might be a faint speech signal $\{s_n\}$ masked by deterministic 'noise' $\{x_n\}$ coming from some kind of vibrating machinery. In both cases, the mathematical problem amounts to separating $\{u_n\}$ into its two components $\{x_n\}$ and $\{s_n\}$. Several schemes exist for performing this task, one of which was recently developed by the Systems Theory Group and which we illustrate here. The basic idea behind this algorithm is to perform the one step prediction as in fig. 6 and examine the difference $d_n = G(u_n, u_{n-1}) - u_{n+1}$. If s_n, s_{n-1} and s_{n+1} were zero, d_n would be negligible, as in fig. 6. On the other hand, if these values of s_n are large, we can typically expect d_n to be large also. This is indeed so, and under certain further assumptions it is possible to estimate the values of $\{s_n\}$ from those of $\{d_n\}$.

As an example of this we took the signal in fig. 7(a), reduced its amplitude by a factor of 10^4 and added it to the $\{x_n\}$ of fig. 2 (corresponding to a signal-to-noise ratio of –80dB!). The result of applying the extraction algorithm is shown in fig. 7(b). Although the extracted signal is still contaminated by some noise, it is clear that its major features have been recovered extremely accurately. Such performance is well beyond the capabilities of any conventional linear signal processing technique such as those based upon auto-regressive and moving average (ARMA) models. Intriguingly, the performance increases as the signal-to-noise ratio decreases (down to some lower limit imposed by numerical inaccuracy), in complete contrast to traditional time series methods.

7. APPLICATIONS

In the preceding sections we have suggested that what appears to be random noise may sometimes be caused (at least in part) by deterministic chaos. We have also shown how (in simple cases) one can make use of the ability to predict chaotic sequences on short timescales to achieve really rather remarkable levels of performance in terms of 'noise' suppression (or equivalently, signal extraction). Although this whole subject has created considerable excitement in the academic world, where it has proved its worth in the analy-

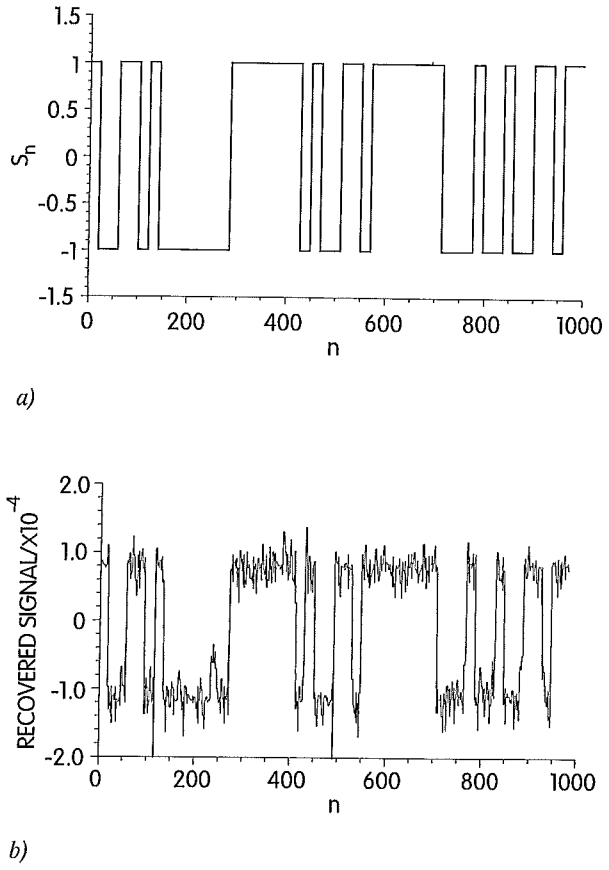


Fig. 7. a) Trial signal used for extraction experiment, and b) extracted signal

sis of time series in rather contrived experimental situations, it is worth concluding by highlighting its potential usefulness (and limitations) in more practical applications.

A number of possibilities immediately suggest themselves and doubtless the reader can think of others. These include mechanical machinery (including turbines, generators and similar rotors) where one might be interested in vibration monitoring or fault detection (such as crack formation which leads to a subtle change in the observed time series.). There may be instances in communication systems where some of the 'noise' may be removed, enhancing the detection of weak signals (there is some evidence that multi-path propagation has chaotic dynamics). Similarly, in radar and sonar there may be specific instances where signal-to-noise ratios can be improved. Each application will require careful optimization of these new techniques and, ideally, their integration alongside existing methods (such as fast Fourier transforms (FFTs) and ARMA models, which are obviously not going to be replaced, only complemented). Unlike FFTs (which work in the frequency domain), the new methods work in the temporal domain, which makes them particularly suited to handling transient or localized phenomena. Finally, we stress again that the radial basis function approach can be applied in principle to the modelling of any non-linear process (not just chaos), and thus may prove useful in applications such as tracking for example.

We have already hinted that computational requirements mean that one has little hope in dealing with systems where the embedding dimension m is greater than about 10. In such cases one probably has to resort to conventional methods. There also remain many questions as to how these methods work when the signal is a mixture of low-dimensional chaos and higher-dimensional 'noise', or when the signal to be extracted is not slowly varying with respect to the background fluctuations. Much work also remains to be done on combining these new techniques with conventional (ARMA) methods so as to gain the best of both worlds. Despite these limitations, it is clear that the subject has a promising future and will develop rapidly once more realistic problems are tackled.

BIBLIOGRAPHY

An excellent introduction to chaos and the mathematics of chaotic dynamics may be found in:

SCHUSTER, H., 'Deterministic Chaos: an Introduction', VCH Publishers, Weinheim, 1989.

For an expanded discussion of many of the ideas presented in this paper, we recommend the following:

BROOMHEAD, D. S. and LOWE, D., 'Multivariable functional interpolation and adaptive networks', Complex Systems, *2*, p. 321-355, 1988.

CASDAGLI, M., 'Nonlinear prediction of chaotic time series', Physica D, *35*, p. 335-356, 1989.

CASDAGLI, M., EUBANK, S., FARMER, J. D. and GIBSON, J. 'State space reconstruction in the presence of noise', Physica D, *51*, p. 52-98, 1991.

FARMER, J. D. and SIDOROWICH, J. J., 'Predicting chaotic time series', Phys. Rev. Lett., *59*, p. 845-848, 1987.

FARMER, J. D. and SIDOROWICH, J. J., 'Optimal shadowing and noise reduction', Physica D, *47*, p. 373-392, 1991.

GRASSBERGER, P., SCHREIBER, T. and SCHAFFRATH, C., 'Nonlinear time sequence analysis', International J. of Bifurcation and Chaos, *1*, p. 521-547, 1992.

STARK, J. and ARUMUGAM, B.V., 'Extracting slowly varying signals from a chaotic background', International J. of Bifurcation and Chaos, to appear June 1992.