Research

Estimating the distribution of dynamic invariants: illustrated with an application to human photo-plethysmographic time series

Michael Small*

Address: Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Email: Michael Small* - ensmall@polyu.edu.hk

* Corresponding author

Abstract

Dynamic invariants are often estimated from experimental time series with the aim of differentiating between different physical states in the underlying system. The most popular schemes for estimating dynamic invariants are capable of estimating confidence intervals, however, such confidence intervals do not reflect variability in the underlying dynamics. We propose a surrogate based method to estimate the expected distribution of values under the null hypothesis that the underlying deterministic dynamics are stationary. We demonstrate the application of this method by considering four recordings of human pulse waveforms in differing physiological states and show that correlation dimension and entropy are insufficient to differentiate between these states. In contrast, algorithmic complexity can clearly differentiate between all four rhythms.

1. Background

Various dynamic invariants are often estimated from time series in a wide variety of scientific disciplines. It has long been known that these estimates (and in particular correlation dimension estimates) alone are not sufficient to differentiate between chaos and noise. Most notably, the method of surrogate data [1] was introduced in an attempt to reduce the rate of false positives during the hunt for physical examples of chaotic dynamics. Although it is not possible to find conclusive evidence of chaos through estimation of dynamic invariants, surrogate methods are often used to generate a distribution of statistic values (i.e. the estimates of the dynamic invariant) under the hypothesis of linear noise. In the most general form, the standard surrogate methods can generate the distribution of statistic values under the null hypothesis of a static monotonic nonlinear transformation of linearly filtered noise.

In this communication, we introduce a significant generalisation of a recent surrogate generation algorithm [2,3]. The *pseudo-periodic surrogate* (PPS) algorithm allows one to generate data consistent with the null hypothesis of a noise driven periodic orbit – provided the data exhibits pseudo-periodic dynamics. Previously, this algorithm has been applied to differentiate between a noisy limit cycle, and deterministic chaos. By modifying this algorithm and applying it to noisy time series data, we are able to generate surrogate time series that are independent trajectories of the same deterministic system, measured via the same imperfect observation function. That is, we assume that there is a deterministic dynamical system subject to additive independent and identically distributed (i.i.d.) observational noise. This ensemble of *attractor trajectory surrogates* (ATS) can then be used to estimate the distribution of statistic values for estimates of any statistic derived from these time series.

The statistics of greatest interest to us are dynamic invariants of the underlying attractor, and in particular correlation dimension and entropy estimates provided by the *Gaussian kernel algorithm* (GKA) [4,5]. Our choice of the GKA is entirely arbitrary, but based on our familiarity with this particular algorithm. True estimation of dynamic invariants from noisy data is a process fraught with difficulty, in this paper we are only concerned with estimating the distribution of estimates. To emphasise this point further we repeat out analysis with another quantity, Lempel-Ziv complexity [6], which *does not* constitute a dynamics invariant. Nonetheless, our algorithm provides a reliable estimate of the distribution of statistic values for this statistic as well.

An important application for the ATS technique is to determine whether dynamic invariants estimated from distinct time series are significantly different. The question this technique can address is whether (for example) a correlation dimension of 2.3 measured during normal electrocardiogram activity is really distinct from the correlation dimension of 2.4 measured during an episode of ventricular tachycardia [7,8]. Estimates of dynamic invariants (including the GKA [4,5]) often do come with confidence intervals. But these confidence intervals are only based on uncertainty in the least-mean-square fit, not the underlying dynamics. Conversely, it is standard practice to obtain a large number of representative time series for each (supposedly distinct) physical state, and compare the distribution of statistic values derived from these. But, this approach is not always feasible: in [7,8] for example, the problem is not merely that these physiological states are difficult and dangerous to replicate, but that inter-patient variability makes doing so infeasible.

In the remainder of this paper we describe the new ATS algorithm and demonstrate that it can be used to estimate the distribution of dynamic invariant estimates from a single time series of a known dynamical system (we demonstrate this with the Hénon map and the chaotic Rössler system). We then apply this same method to four recordings of human pulse waveforms, measured via photoplethysmography [9,10]. Each of the four recordings correspond to a distinct physiological state. We compute correlation dimension and entropy using the GKA method and show that the expected distribution of correlation dimension and entropy estimates are insufficient to differentiate between these four physiological states. In contrast, we show that algorithmic complexity can clearly differentiate between all four rhythms.

In Section 2 we describe the algorithm we employ in this paper, and in Section 2.2 we demonstrate that, for suitable parameter values, this technique will preserve the deterministic dynamics of the underlying system. In Section 3 we present some numerical case studies and in Section 4 we finally present our conclusions.

2. Attractor trajectory surrogates
In the first part of this section we will review the PPS algorithm presented in [2] and describe the novel features of the ATS approach. In section 2.2 we examine the foundation of this technique's ability to preserve the underlying deterministic dynamics.

2.1 The algorithm
In what follows we assume that the measured scalar time series x_t represents discretely sampled measurements of a deterministic dynamical system (possibly continuous) under the influence of observational noise. In other words, the dynamics are determined by a smooth manifold M and deterministic evolution operator $\phi : M \to M$. The output of the evolution of an initial condition $m_0 \in M$ under ϕ (i.e. $m_t = \phi^t(m_0)$) are observed via the differentiable function $h : M \to \mathbf{R}$. Unfortuantely, experimental measurement is not perfect and the observed time series $\{x_t\}$ is subject to observational noise, hence, $x_t = h(\phi^t(m_0)) + \varepsilon_t$ where $\varepsilon_t \sim \mathcal{N}$ is drawn from some stationary noise distribution $\mathcal{N}$. For the case of dynamic noise, the situation is complicated further as the evolution of m_t is governed by $m_{t+1} = \phi(m_t) + \xi_t$ where ξ_t is stochastic.

The ATS algorithm may now be described as follows. Embed the observed scalar time series $\{x_t\}$ to obtain a vector time series $\{z_t\}$, $z_t \in \mathbf{R}^d$, of N observations. The choice of embedding is arbitrary, but has been adequately discussed in the literature (there are numerous works in this field, [11] for example, provides references to several of them). We assume that the embedding is such that there exists a continuously differentiable map $\Xi : M \to \mathbf{R}^d$ between the underlying manifold M and the embedding space $\mathbf{R}^d$ such that both Ξ and $D\Xi$ are one-to-one. Under these conditions, the dynamics of (ϕ, M) and the evolution of $z_t = \Xi(m_t) \in \mathbf{R}^d$ are considered to be equivalent. From the embedded time series, the surrogate is obtained as follows. Choose an initial condition, $w_1 \in \{z_j | j = 1, ..., N\}$. Then, at each step i, choose the successor to w_i with probability

$$P(w_{i+1} = z_{j+1}) \propto \exp \frac{-\left\| w_i - z_j \right\|}{\rho} \qquad (1)$$

where the *noise radius* ρ is an as-yet unspecified constant. That is, the successor of w_i, w_{i+1} is chosen to be the point

z_{j+1} with probability proportional to $\exp\dfrac{-\left\|w_i - z_j\right\|}{\rho}$,

where z_j is the antecedent of z_{j+1}. In other words, the successor to w_i is the successor of a randomly chosen neighbour of w_i. Equation (1) may then be written as

$$P(w_{i+1} = z_{j+1}) = \frac{P_{i,j}}{\sum_{k=1}^{N} P_{i,k}}$$

where $P_{i,k} = \exp\dfrac{-\left\|w_i - z_j\right\|}{\rho}$ (and similarly,

$P_{i,j} = \exp\dfrac{-\left\|w_i - z_j\right\|}{\rho}$). Finally, from the vector time series

$\{w_i\}$ the ATS $\{s_i\}$ is obtained by projecting w_i onto $[1\ 0\ 0\ 0 \cup 0] \in \mathbf{R}^d$ (the first coordinate). Hence

$$s_t = w_t \cdot [1\ 0\ 0\ 0 \cup 0] \qquad (2)$$

In [2,3] this algorithm was shown to be capable of differentiating between deterministic chaos and a noisy periodic orbit. In the context of the current communication we assume that $\{x_t\}$ is contaminated by additive (but possibly dynamic) noise and we choose the noise radius ρ such that the observed noise is replaced by an independent realisation of the same noise process. Furthermore, we assume that the deterministic dynamics are preserved by suitable choice of embedding parameters. Under these two assumptions, $\{z_t\}$ and $\{w_t\}$ have the same invariant density and $\{x_t\}$ and $\{s_t\}$ are therefore (noisy) realisation of the same dynamical system with (for suitable choice of ρ) the same noise distribution. We illustrate this more precisely in the following section.

2.2 Invariance
As in [2,3] the problem remains the correct choice of ρ. This is the major difference between the ATS described here and the PPS of [2,3]. However, since the null hypothesis we wish to address is different from (and more general than) that of the PPS, choice of ρ for the ATS is less restrictive. For $t = T$ given, one can compute $P(w_{t+1} \neq z_{i+1} \wedge ||w_t - z_i|| = 0|t = T)$ directly from the data by applying (1) to the embedded time series(we use the symbol $\wedge$ here in the usual manner to denote logical conjunction). Assuming the process is ergodic (that is, ergodic with respect to the standard measure – this assumption is sufficient rather than necessary) one can then sum

$$P(w_{i+1} \neq z_{j+1} \wedge \left\|w_i - z_j\right\| = 0) =$$
$$\frac{1}{N}\sum_{T=1}^{N} P(w_{i+1} \neq z_{j+1} \wedge \left\|w_i - z_j\right\| = 0 \Big| j = T)$$
$$(3)$$

to get the probability of a temporal discontinuity in the surrogate at any time instant. By temporal discontinuity we mean that $w_i = z_j$ but $w_{i+1} \neq z_{j+1}$. That is, a point where the surrogate trajectory does not exactly follow the data. There is a one-to-one correspondence between a value $p = P(w_{t+1} \neq z_{i+1} \wedge ||w_t - z_i|| = 0)$ and ρ, and we choose to implement (1) for a particular value of p (i.e. a particular transition probability) rather than a specific noise level ρ. In what follows we find that studying intermediate values of p ($p \sim 0.1$) is sufficient. For $p \in [0.1, 0.8]$ the qualitative behaviour over the corresponding narrow range of ρ is uniform. We choose to illustrate with $p = 0.1$, but the results for other choices are similar. Of course, for $p \rightarrow 1$ or $p \rightarrow 0$ the algorithm will not work well.

Now, suppose that the embedding parameters τ and d_e have been selected correctly and the noise in the data is not too large, then the transformation $x_t \mapsto z_t$ dictated by these parameters is an embedding. That is, the operator $\Xi : M \rightarrow \mathbf{R}^d$ with $\Xi(m_t) = z_t$ (in the absence of noise) and its derivative $D\Xi$ are both one-to-one. Hence, the dynamic evolution of $z_j \in \mathbf{R}^{d_e}$ can be represented by

$$z_{j+1} = \Phi(z_j) + e_j \qquad (4)$$

where $\Phi(\cdot)$ is diffeomorphic to the true evolution operator (i.e. $\Phi = \Xi \circ \phi \circ \Xi^{-1}$ where $\phi : M \rightarrow M$ is the underlying evolution operator, defined earlier) and e_j are uncorrelated noise vectors (corresponding to the terms ε_t and possibly ξ_t described earlier). Now we consider the process of constructing a surrogate. Let $\left\{w_i\right\}_{i=1}^{N}$ denote the surrogate vector time series of length N. Clearly, setting $w_1 = z_k$ for some randomly chosen k is simply some new initial condition. Now, $w_{i+1} = z_{j+1}$ where j is chosen randomly from a distribution such that $||w_i - z_j||$ is small. Let $\varepsilon_i = z_j - w_i$ corresponding to the small (random) perturbation introduced by selection according to (1), then

$$w_{i+1} = \Phi(w_i + \varepsilon_i) + e_j. \qquad (5)$$

Note that, ε_j is the perturbation introduced in taking z_j's successor to be the successor of w_i (it is a *dynamic noise*

term, and it is a perturbation introduced by the ATS method). Conversely, e_j is the dynamic error in applying Φ (this term is inherent to the data, and to our model of the data). By taking n-th iterates of (4) and (5) we see that the two noise terms e_j and ε_{i+1} will combine. In other words, from (5) we get

$$w_{i+2} = \Phi(\Phi(w_i + \varepsilon_i) + e_j + \varepsilon_{i+1}) + e_{j+1}, \qquad (6)$$

and so on. Suppose that $e_j \sim \mathcal{D}$ where $\mathcal{D}$ is some noise distribution. Then, for the surrogates $\{s_t\}$ to be a new realisation of the system that generated $\{x_t\}$ we require that $e_j + \varepsilon_{i+1} \sim \mathcal{D}$. But this is equivalent to the condition that $z_j - w_i \sim k\mathcal{D}$ for sufficiently small k. Hence, the critical issue is the choice of ρ such that the these two noise terms are drawn from the same distribution and that therefore the surrogate dynamic (5) is equivalent to (4). This requires sufficient data, ergodicity, and ρ small enough. Note that, as ρ becomes smaller and the surrogate data become more like realisations of the same system, we also see less randomisation. This is a natural and unavoidable tradeoff.

3. Results

The following subsections present the application of this method for data generated from the Hénon map (section 3.1), the Rössler system (section 3.2) and experimental measurements of human pulse pressure waves (section 3.3).

3.1 The Hénon map

One potential difficulty of this method is that the stretching-and-folding characteristic of Smale horseshoe type chaos could easily destroy the dynamics of (5) and therefore produce surrogate trajectories that short-cut across the attractor. Although we can see from equations (5) and (6) that for sufficiently small perturbations this will never be the case, we would like to test this possibility in practise. For this purpose we apply the method described in the previous section to the extremely well studied Hénon map: one of the archetypes of Smale horseshoe chaos.

Figure 1 illustrate typical ATS calculations for this data set. Using short (1000 point) sections of the Hénon system, with the addition of observational noise (the Figures show 1% and 10% noise levels), we computed typical ATS data for different values of transition probability p. We find that in almost all cases (see Figure 1) the results for the ATS data agree qualitatively with the data. Comparison of estimated dynamic invariant (results omitted) confirm this. In all cases, for moderate range of p (i.e. p neither approaching 0 or 1) and moderate observational noise, we find data and surrogate agree closely. When this same computation was repeated for *dynamic* noise, we

found data and surrogates to be similarly indistinguishable (see Figure 2): except for the case of large p and small noise (in this case, 1% dynamic noise and $p = 0.8$). Note that, for the Hénon map larger values of dynamic noise will actually force the system into an unstable régime.

3.2 The Rössler system

We now demonstrate the applicability of this method for a more realistic example: noisy time series data simulated from the Rössler differential equations (during "broadband" chaos). We integrated (one thousand points with a time step of 0.2) the Rössler equations both with and without multidimensional dynamic noise at 5% of the standard deviation of the data. As far as possible, we generated realisations of the Rössler system that superficially resemble the physiological data of 3.3. The purpose of this is to provide a more realistic test of our method. We then studied the x-component after the addition of 5% observational noise. We selected embedding parameters using the standard methods (yielding $d_e = 3$ and $\tau = 8$) and then computed ATS surrogates for various exchange probabilities $p = 0.05, 0.1, 0.15, ..., 0.95$.

For the data set and each ensemble of surrogates we then estimated correlation dimension D, entropy K and noise level S using the GKA algorithm [4,5]. The GKA embedding used embedding dimension $m = 2, 3, ..., 10$ and embedding lag of 1. It is important to note that, a correlation dimension estimate is not the same thing as the actual correlation dimension. In particular, this algorithm estimates correlation dimension and noise level simultaneously (as well as entropy). A lower correlation dimension (associated with the presumed determinism in the system) is accompanied by an increase in the estimated noise level. That is, the estimated dimension can be lower because the algorithm is attributing more of the variation in the data to noise, and therefore estimating a higher noise level (and hence, in some case, the correlation dimension falls below 1). Similarly, the fact that the entropy is negative in the first case is associated with the system noise. Nonetheless, we are using these numbers only as measures, that is, as test statistics. Figure 3 depicts the results when the GKA is applied with embedding dimension $m = 4$ and the exchange probability is $p = 0.1$. Other values of m gave equivalent results, as did various values of p in the range [0.1, 0.8].

For such moderate p we found that the estimate of noise S from the GKA algorithm coincided for data and surrogates, but this was often not the case for more extreme values of p. This estimate of signal noise content is therefore a strong test of the accuracy of the dynamics reproduced by the ATS time series. One expects this to be the case as noise level is precisely the parameter upon which the ATS method depends. Furthermore to confirm the spread of
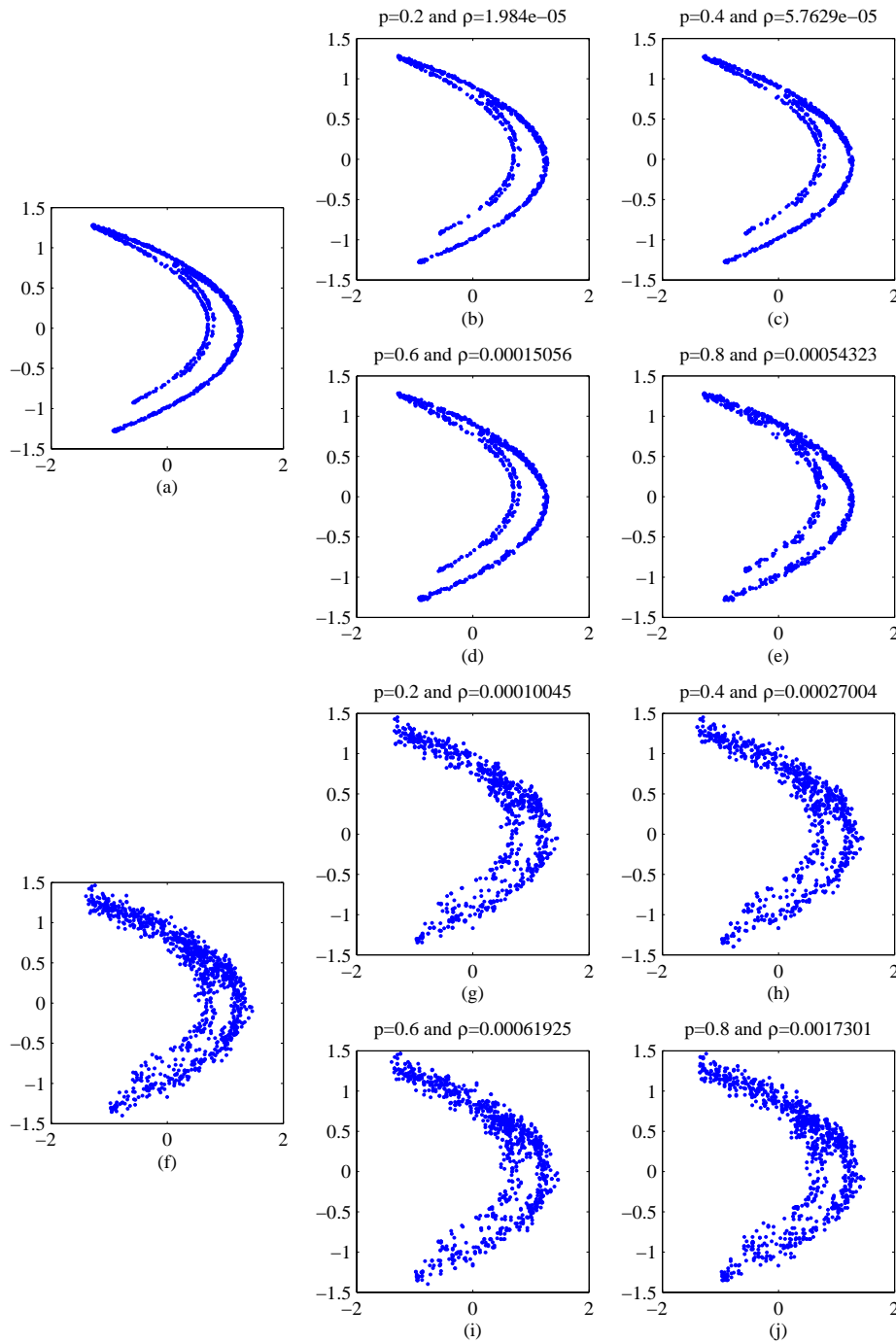
Figure 1
Sample reconstructed attractors for data and surrogates of the Hénon map. Panels (a) and (f) are embedded time series data from the *x*-component of the Hénon system with the addition of 1% and 10% observational noise (respectively). The remaining panels are representative ATS time series. Panels (b), (c), (d) and (e) are surrogates for panel (a), and Panels (g), (h), (i) and (j) are for panel (f). Each surrogate is computed with a different level of transition probability *P*. In panels (b) and (g), $p = 0.2$; in panels (c) and (h), $p = 0.4$; in panels (d) and (i), $p = 0.6$; and, in panels (e) and (j), $p = 0.8$. In each case the attractors reconstructed from the surrogates have the same qualitative features as that of the data – with the possible exception of panel (e). The likely reason for this noted exception is the relatively high transition probability ($p = 0.8$) and the relatively low noise level (1%). Of course, for smaller values of *p* (i.e. *p* = 0.1) the similarity is even more striking.
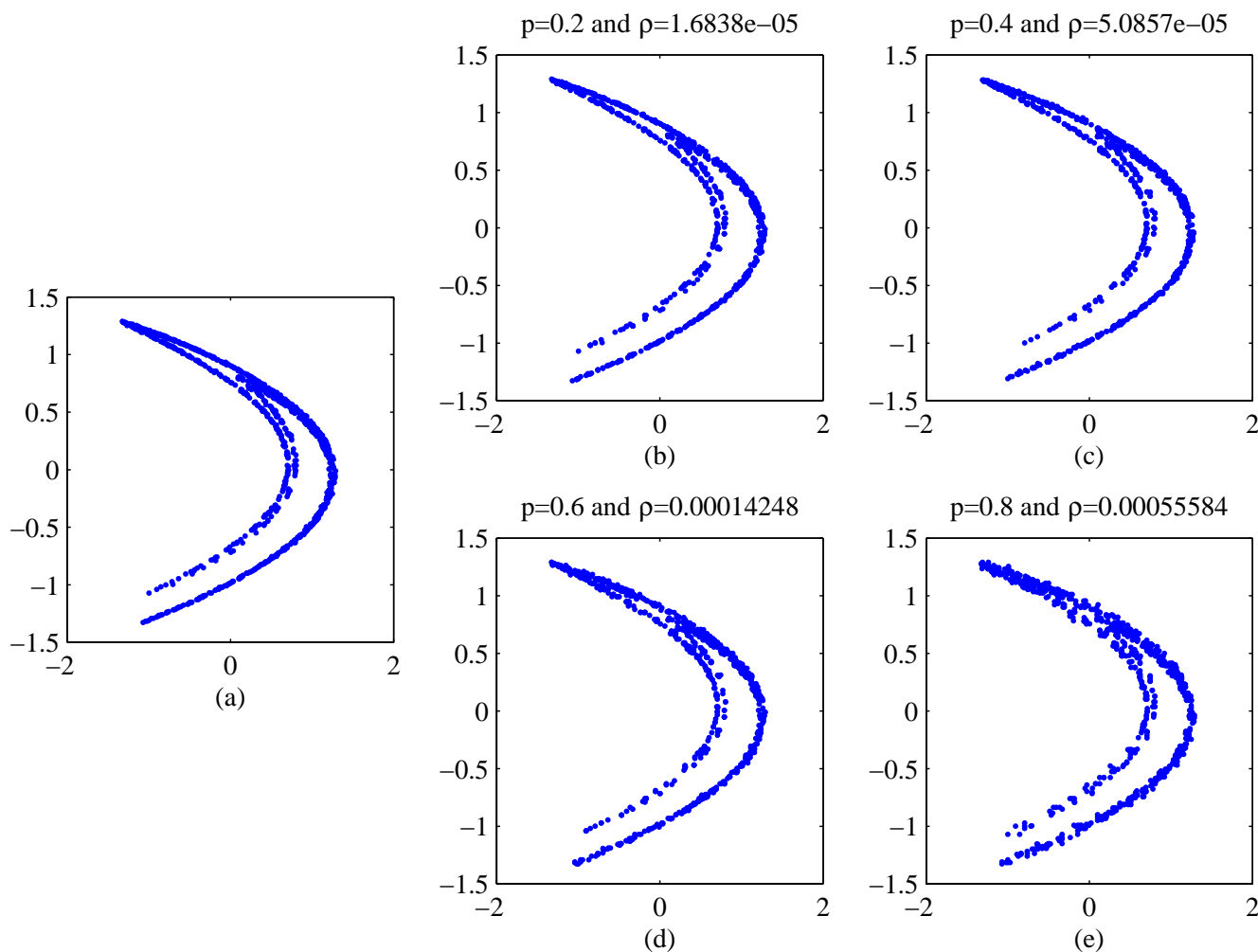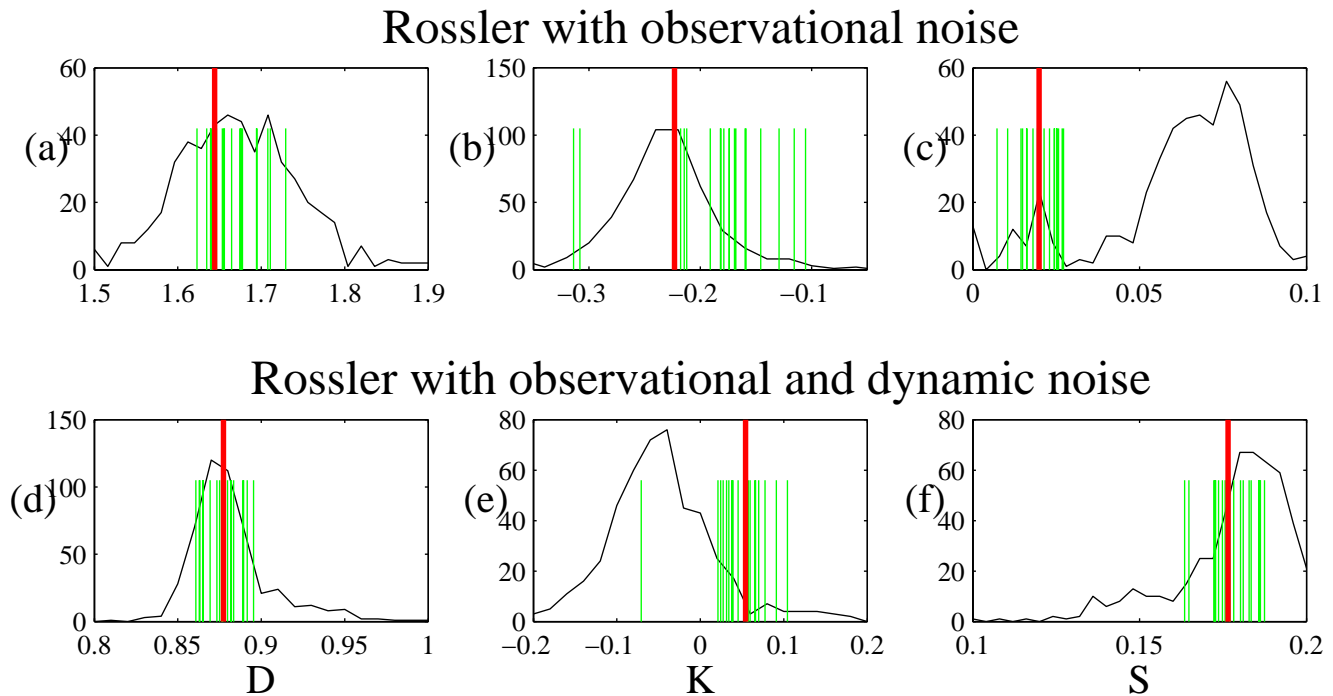
Figure 2
Sample reconstructed attractors for data andsurrogates of the Hénon map. Panel (a) is an embedded time series
data from the *x*-component of the Hénon system with the addition of 1% *dynamic* noise. The remaining panels are representa-
tive ATS time series. Each surrogate is computed with a different level of transition probability *P*. In panel (b) *p* = 0.2; in panel
(c) *p* = 0.4; in panel (d) *p* = 0.6; and, in panels (e) *p* = 0.8. In each case the attractors reconstructed from the surrogates have
the same qualitative features as that of the data.

the data we also estimated *D*, *K*, and *S* for 20 further real-
isations of the same Rössler system (with different initial
conditions). In each case, as expected, the range of these
values lies within the range predicted by the ATS scheme.
We do see, for example, in Figure 3(c) that the range of
noise level exhibited by the true Rössler system is not as
expansive as that for the surrogates (to some extent, we
can also observe the same problem with entropy in Figure
3(e)). This is due to the fact that the ATS method can be
made to introduce more randomisation than absolutely
necessary. By tuning down the randomisation we (obvi-
ously) will converge to the true data. By increasing the
randomisation we cover an ever widening range, which
will always include the true value. For large randomisa-

tion, and for statistics that are most sensitive to noise (in
this case *K* and *S*) there may also be some bias – the
observed difference in the means. Although it is desirable
that both distributions coincide exactly, it is re-assuring
(and sufficient) that the ATS distribution contains the true
distribution.

3.3 Photo-plethysmographic recordings
We now consider the application of this method to photo-
plethysmographic recordings of human pulse dynamics
over a short time period (about 16.3 seconds). We have
access to only a limited amount of data representative of
each of four different dynamic regimes. In any case, we
would expect the system dynamics to change if measured

Figure 3
Distribution of statistics *D*, *K* and *S* for short and noisy realisations of the Rössler system. The histogram shows the distribution of statistic estimates (*D*, *K* and *S*) for 500 ATS time series generated from a 1000 point realisation of the Rössler system. The tall vertical line on each plot is the comparable value for the data and the shorter vertical lines indicate 20 independent realisations of the same process. The top row of figures depicts results for the Rössler system with observational noise only, the bottom row of figures has both observational and dynamic noise. Panels (a) and (d) show correlation dimension estimates, (b) and (e) are entropy, and (c) and (f) are noise level.

over a significantly longer time frame. The data collection and processing with the methods of nonlinear time series analysis are described in [9,10]. Previously, we have studied nonlinear determinism in cardiac dynamics measured with electrocardiogram (ECG) [7,8]. Although we do not consider ECG data here, this data would be another useful system to examine with these methods. Actually, the problem with ECG data is that we have too much data and it is therefore difficult to fairly select a "representative" small number of short time series. However, we intend to examine this data more carefully in the future. However, we do note in passing that both PPG and ECG are measures of cardiac activity and are therefore potentially equivalent [12,13]. The four data sets we examine in this communication are depicted in Figure 4.

For each data set we repeated the analysis described for the Rössler time series. Results for GKA embedding dimension *m* = 6 and *p* = 0.1 are depicted in Figure 5. As with the Rössler system, variation of the parameters *m* and *p* did not significantly change the results. We find that in every

case (except for extreme values of *p*) the distribution of *D*, *K* and *S* estimated from the ATS data using the GKA included the true value. Most significantly, this indicates that the range of values of *p* is appropriate. Moreover, these results are consistent with the hypotheses that the noise is effectively additive and can be modelled with this simple scheme, and that the underlying deterministic dynamics can be approximated with a local constant modelling scheme.

We also estimated the statistics *D*, *K* and *S* for additional available data (subsequent, contiguous, but non-overlapping) from each of the four rhythms. This small amount of data afforded us two or three additional estimates of each statistic for each rhythm. For the unstable and quasi-stable rhythm we observed good agreement. For the stable (normal and post-operative) rhythms, this is not the case. On examination of the data we find that this result is to be expected. Both the stable rhythms undergo a change in amplitude and baseline subsequent to the end of the original 16 second recording, this non-stationarity is reflected
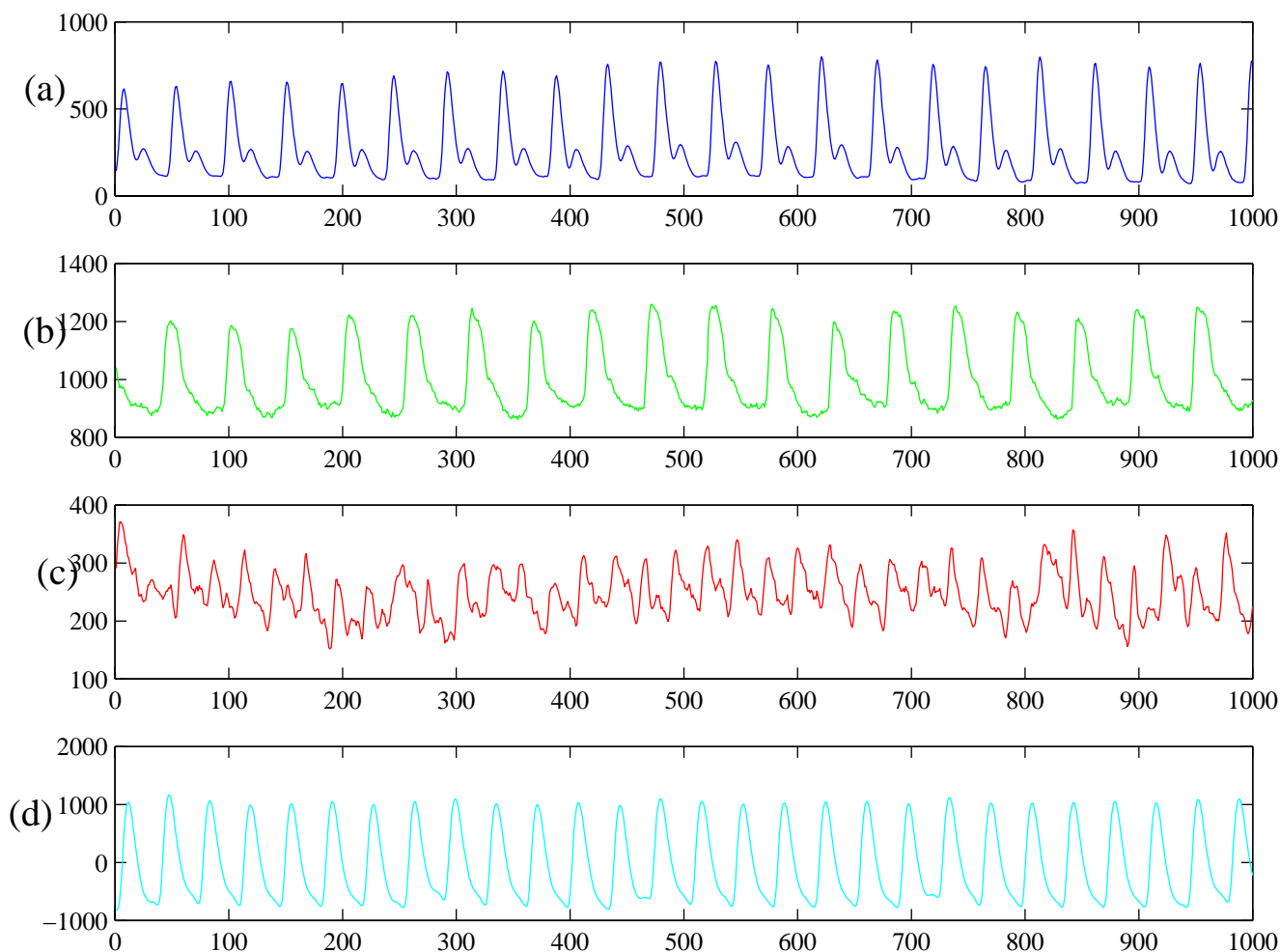
Figure 4
Human pulse waveform recorded withphoto-plethysmography. Four recordings of human pulse waveform (61 Hz) in four different physiological conditions. The four time series correspond to: (a) normal, (b) quasi-stable, (c) unstable, and (d) post-operative (stable).

in the results. This same non-stationarity has also been observed independently in Bhattacharya and co-workers [9,10].

We now return to the question that the ATS test was designed to address: can we differentiate between these four rhythms based on the GKA? Figure 6 provides the answer. In Figure 6 we see the estimated distribution of statistic values (*D*, *K* and *S*) for each of the four rhythms shown in figure 4. Clearly (and not surprisingly), the correlation dimension estimate and noise level of the unstable rhythm is significantly different from the other three rhythms.

Our analysis indicates that, contrary to what one may expect from individual measurements, the stable or

"quasi-stable" rhythms cannot be properly distinguished based on these nonlinear statistics derived from the GKA. Moreover, we find that entropy estimated with the GKA algorithm *K* is of no use in differentiating between any of these four rhythms. Although it is not the purpose of this paper to provide a discriminating statistic for this data, it would be nice to do so. Therefore, in Figure 7 we repeat the calculation of surrogates and statistic distribution for the same data, but using algorithmic complexity (see [11] and the references therein) with binary, ternary, and quaternary encodings with equal likelihood for each symbol. Using this scheme it can be seen from Figure 7 that it is possible to distinguish, with a high level of certainty between three of these rhythms. Distinguishing between all four is also possible, with a small likelihood of error (see Figure 7(a)).
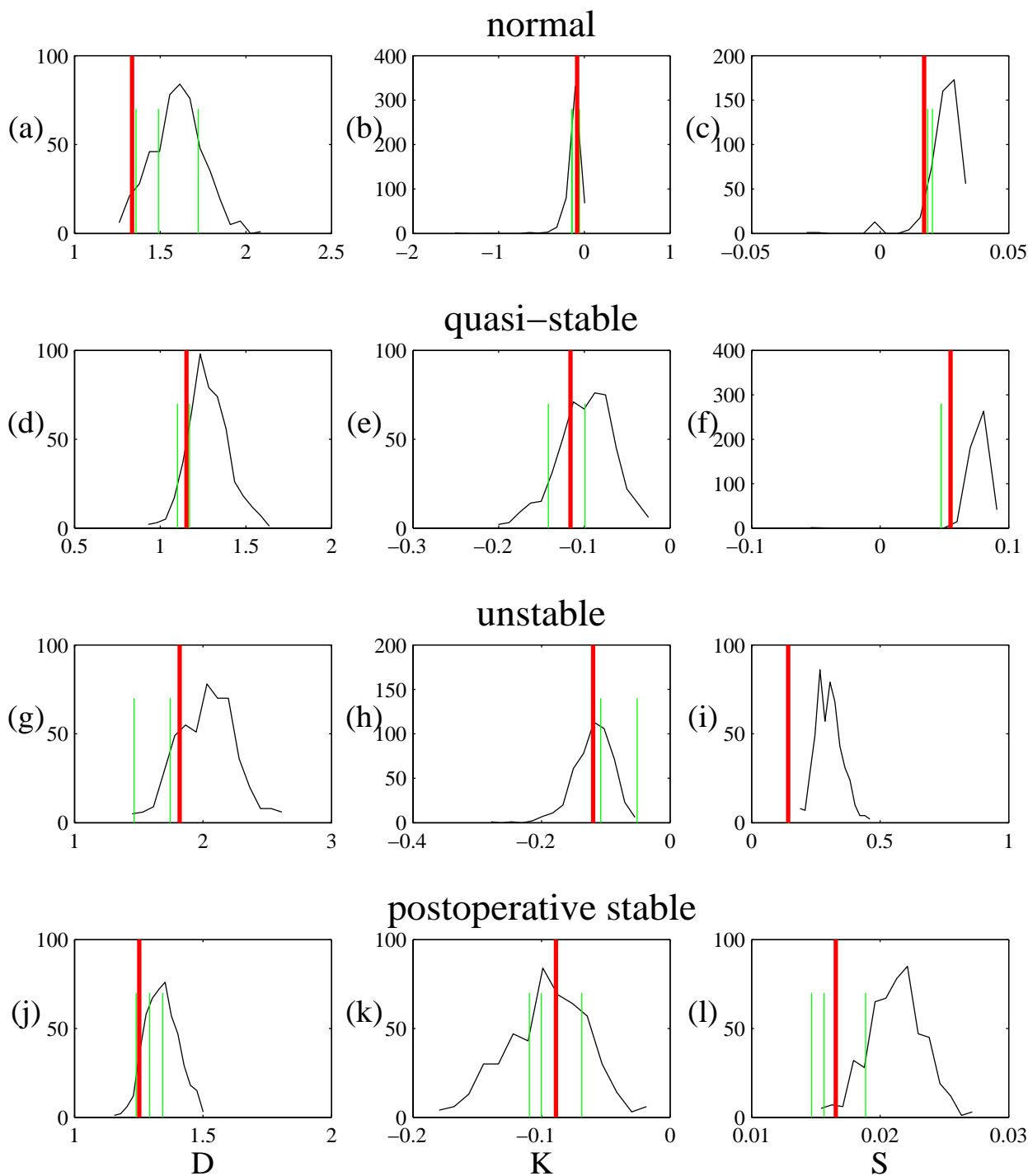
Figure 5
Distribution of statistics *D*, *K* and *S* for human pulse waveforms. The histogram shows the distribution of statistic esti-
mates (*D*, *K* and *S*) for 500 ATS time series generated from each of the four time series depicted in figure 4. The taller vertical
line on each plot is the comparable value for the shorter vertical lines are for the (limited) subsequent data recorded from each
patient. In each case only two or three subsequent contiguous but non-overlapping time series were available. The figures are:
(a) correlation dimension (*D*), (b) entropy (*K*), and (c) noise (*S*) for the normal rhythm; (d) *D*, (e) *K*, and (f) *S* for the quasi-stable
rhythm; (g) *D*, (h) *K*, and (i) *S* for the unstable rhythm; and (j) *D*, (k) *K*, and (l) *S* for the post-operative stable rhythm.
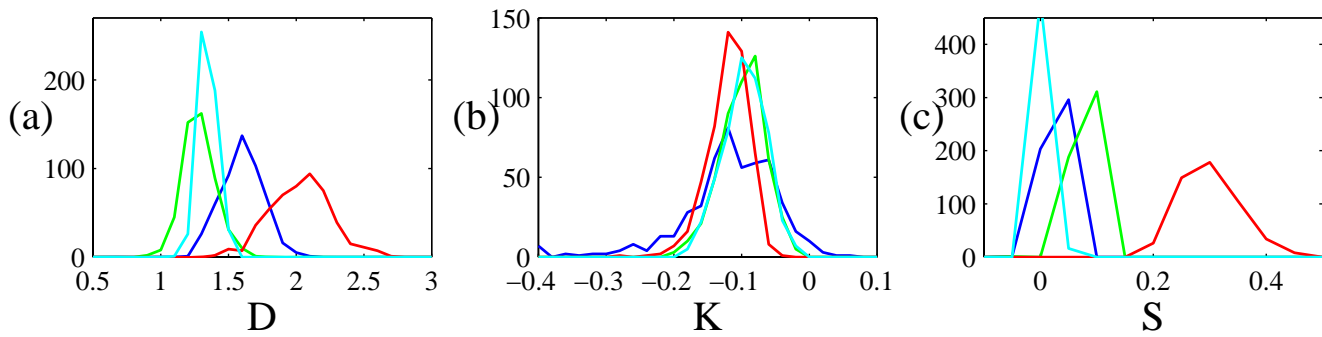
Figure 6
Discriminating power of the statistics *D*, *K* and *S* for human pulse waveforms. The distribution (a binned histogram) of statistic values estimated via the ATS method (as described in figure 5) for each of the four distinct physiological waveforms is shown. The four rhythms correspond to those in Figure 4. These figures show that correlation dimension alone is sufficient to differentiate between three of these four physiological states: on the left, "post-operative" and "quasi-stable" are indistinguishable, the correlation dimension for "normal" is bigger, and "unstable" is larger again We see that these three statistics are insufficient to differentiate between the "quasi-stable" and "post-operative" states, moreover, there is considerable overlap with the "normal" group.

4. Conclusion

The results of this analysis are in general agreement with those presented in [9,10]. Independent linear surrogate analysis [1] has confirmed that each of these four rhythms is inconsistent with a monotonic nonlinear transformation of linearly filtered noise (these calculations are routine, and not presented in this paper). The only significant difference is that the correlation dimension estimates we present here are significantly lower than those in [9,10]. This is due to the different correlation dimension algorithm. Unlike the algorithm employed in [9,10], the GKA seperates the data into purely deterministic and stochastic

components, and hence estimates both *D* and *S*. The correlation dimension estimated in [9,10] is the combined effect of both components of the GKA.

Although we have considered the specific application to human pulse dynamics, the algorithm we have proposed may be applied to a wide variety of problems. We have shown that provided time delay embedding parameters can be estimated adequately, and an appropriate value of the exchange probability is chosen, the ATS algorithm generates independent trajectories from the same dynamical system. When applied to data from the Rössler system
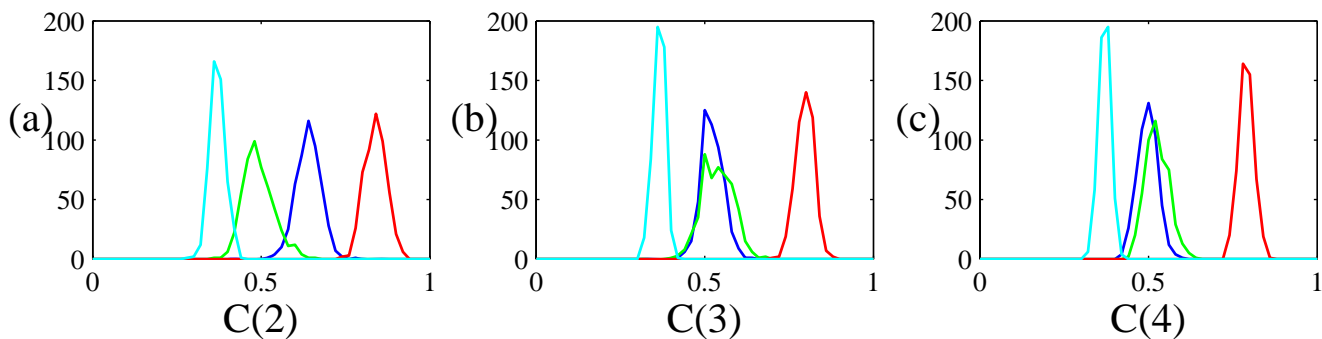

Figure 7
Discriminating power of complexity for human pulse waveforms. The distribution (a binned histogram) of statistic values estimated via the ATS method (as described in figure 5) for each of the four distinct physiological waveforms is shown. The four rhythms correspond to those in figure 4. These figures show that complexity with 2, 3, and 4, symbols (plots (a), (b), and (c), respectively) is sufficient to differentiate between at least three of these four physiological states. The lowest complexity corresponds to "post-operative" state, the next highest to "quasi-stable" followed by "healthy" and finally "unstable". As in figure 6 there is considerable overlap between the "normal" and "quasi-stable" samples. However, for complexity with a binary partition (panel (a)) the four rhythms do appear to be distinct.

we confirm this result, and we demonstrate its application to experimental data.

When the ATS algorithm is applied to generate independent realisation of a hypothesis test, one is able to construct a test for non-stationarity. If two data sets do not fit the same distribution of ATS data then they can not be said to be from the same deterministic dynamical system. Unfortunately, the converse is not always true and the power of the test depends on the choice of statistic. The utility of this technique as a test for stationarity remains uncertain.

Acknowledgements

References

1. Theiler J, Eubank S, Longtin A, Galdrikian B, Farmer JD: **Testing for nonlinearity in time series: The method of surrogate data.** *Physica D* 1992, **58:**77-94.
2. Small M, Yu D, Harrison RG: **Surrogate test for pseudo-periodic time series data.** *Physical Review Letters* 2001, **87:**188101.
3. Small M, Tse C: **Applying the method of surrogate data to cyclic time series.** *Physica D* 2002, **164:**187-201.
4. Diks C: **Estimating invariants of noisy attractors.** *Physical Review E* 1996, **53:**R4263-R4266.
5. Yu D, Small M, Harrison RG, Diks C: **Efficient implementation of the Gaussian kernel algorithm in estimating invariants and noise level fromnoisy time series data.** *Physical Review E* 2000, **61:**3750-3756.
6. Lempel A, Ziv J: **On the complexity of finite sequencess.** *IEEE Trans Inform Theory* 1976, **22:**75-81.
7. Small M, Yu D, Simonotto J, Harrison RG, Grubb N, Fox K: **Uncovering non-linear structure in human ECG recordings.** *Chaos, Solitons and Fractals* 2001, **13:**1755-1762.
8. Small M, Yu D, Grubb N, Simonotto J, Fox K, Harrison RG: **Automatic identification and recording of cardiac arrhythmia.** *Computers in Cardiology* 2000, **27:**355-358.
9. Bhattacharya J, Kanjilal P: **Assessing determinism of photo-plethysmographic signal.** *IEEE Transactions on Systems, Man and Cybernetics A* 1999, **29:**406-410.
10. Bhattacharya J, Kanjilal P, Muralidhar V: **Analysis and characterization of photo-plethysmographic signal.** *IEEE Transactions on Biomedical Engineering* 2001, **48:**5-11.
11. Small M: *Applied Nonlinear Time Series Analysis: Applications in Physics, Physiology and Finance, Volume 52 of Nonlinear Science Series A* Singapore: World Scientific; 2005.
12. Zhao Y, Small M: **Equivalence between 'feeling the pulse' on the human wrist and the pulse pressure wave at fingertip.** *International Journal of Neural Systems* 2005, **15:**277-286.
13. Zhao Y, Small M: **Evidence consistent with deterministic chaos in human cardiac data: surrogate and nonlinear dynamical modeling.** *International Journal of Bifurcation and Chaos* 2007. To appear