

## DETECTION OF CHANGING DYNAMICS IN PHYSIOLOGICAL TIME SERIES

**L. M. Hively and V. A. Protopopescu\***

Oak Ridge National Laboratory  
P. O. Box 2009, Oak Ridge, TN 37831-8066  
[hivelylm@ornl.gov](mailto:hivelylm@ornl.gov); [protopopesva@ornl.gov](mailto:protopopesva@ornl.gov)

### ABSTRACT

We present a robust, model-independent technique for quantifying changes in the dynamics underlying nonlinear time-series data. The changes in the dynamics translate into dissimilarities between invariant distributions which are measured via  $L_1$ -distance and  $\chi^2$  statistics. The performance of the new measures is tested on model data and applied to various clinically evaluated physiological data to detect early signs of transition from normal to abnormal regimes. The results show a clear superiority of the new measures in comparison to traditional nonlinear measures, in terms of robustness, timeliness, and discriminating power.

*Key Words:* Nonlinear time series, Physiological data, Nonlinear measures, Forewarning

### 1. INTRODUCTION

Living systems in either normal or pathologic conditions display a rich variety of dynamical behaviors. These behaviors manifest themselves in signals that can be interpreted at various levels, such as: physical, chemical, physiological, clinical, etc. One of the most important problems encountered in time-series analysis is the appropriate characterization of changes in the system's dynamics. This problem is particularly vexing in physiological systems, which usually are complex, nonlinear, nonstationary, and strongly affected by the environment. In the last two decades, since the advent of chaotic dynamics on the scientific stage ([1, 2, 3, 4] and references therein), there has been strongly reenergized interest in interpreting physiological data within a dynamical system framework. The dynamical approach is motivated by several features that are shared by physiological and complex physical systems, namely: multiple time scales, quasiperiodicity, chaos, and self-organization. It is therefore reasonable to assume that, under certain circumstances, one can use the nonlinear dynamics framework to analyze and interpret physiological time series. This approach is expected to complement traditional medical diagnostics, warning, prevention, and cure, with more precisely quantified assessments.

Henceforth, we shall assume that the brain, the heart, the lung, etc., behave like low-dimensional nonlinear systems whose dynamics may vary between (quasi-)periodic and completely irregular (chaotic) [5, 6]. Thus, global aspects of the dynamics of these organs may be legitimately quantified by traditional nonlinear descriptors such as Lyapunov exponents, Kolmogorov entropy, and correlation dimension. While these descriptors are adequate, in general, for

---

discriminating between clear-cut regular and chaotic dynamics, they are not sufficiently sensitive to distinguish between *slightly different* chaotic regimes, especially when data are limited and/or noisy. Therefore, robust and timely forewarning of abnormal clinical events such as epileptic seizures, cardiac fibrillation, acute respiratory syndrome, etc., has remained an outstanding medical challenge, especially for non-hospitalized patients.

To address this problem, we propose quantification of change in physiological time series by defining and using new measures of dissimilarity. Following standard techniques, our method converts time-serial data to a geometric representation that describes the dynamics of the underlying nonlinear system (e.g., brain, heart, lung, etc.) on its attractor in the corresponding phase space (PS). The frequency and sequence of visitation of various points of the attractor are described by a distribution function (DF), which does not change if the dynamics remain unchanged. If the dynamics change, the attractor and the DF changes as well. To compare a test case DF to the base case DF, we define various distances between the DFs. A large distance implies that the system has departed significantly from the base case and can be interpreted as a forewarning of an impending abnormal event, possibly a seizure, a cardiac fibrillation, or an acute breathing difficulty.

Our method combines several original advances to achieve sensitivity which is at least one order of magnitude larger than that obtained to date by competing methods. First, before constructing the phase space distribution, we remove confounding artifacts (such as eye-blinks for the EEG signal) with a new zero-phase quadratic filter. Artifact removal allows detection of dynamical change from single-channel data, thereby enabling less invasive, possibly ambulatory, non-clinical monitoring. Second, we use differential measures of dissimilarity which preserve a much higher content of dynamical information than the traditional measures, which average out dynamical changes by integration over large amounts of data. Third, our technique is rather general and applies to various types of clinical data.

The paper is organized as follows. In Section 2 we discuss traditional nonlinear measures for time series analysis. Section 3 presents our indicators of dynamical change by comparison of the phase-space distribution functions via dissimilarity measures. In Section 4 we present the results of our method as applied to model data and clinical physiological time series. We summarize the conclusions in Section 5.

## 2. NONLINEAR MEASURES FROM TIME SERIES

Our analysis begins with a process-indicative scalar signal,  $e$ , collected from a dynamical system whose dimensionality and structure are, in principle, unknown. This signal is sampled at equal time intervals,  $\tau$ , starting at time  $t_0$ , yielding the time series  $e_i = e(t_0 + i\tau)$ ,  $i = 1, 2, \dots, N-1$ .

Usually, physiologic data are affected by artifacts, arising from eye blinks, muscle twitches, etc. We remove essentially all of these artifacts with a novel zero-phase quadratic filter [7]. The raw data are fitted to a quadratic equation over a moving window of  $2w + 1$  points with the same number of data points,  $w$ , on either side of a central point. The fitted value at this central point is taken as the best estimate for the artifact signal,  $f_i$ . This filter is applied to the  $N$ -point cutset of

raw data,  $e_i$ , yielding  $N - 2w$  points of artifact data. Subsequent analysis uses only the artifact-filtered data,  $x_i = e_i - f_i$ , which is essentially artifact-free. By using the dynamical process reconstruction technique [1, 8], this time series is used to find the main features of the underlying unknown dynamics, including its dimensionality, topology of the attractor, Lyapunov exponents, etc. The reconstruction proceeds as follows: for a suitably chosen lag  $\lambda$  one forms from the scalar time series a  $d$ -dimensional time-delay vector,  $y(i) = [x_i, x_{i+\lambda}, \dots, x_{i+(d-1)\lambda}]$ , for a system with dimensionality  $d$ . The dimensionality  $d$ , which is unknown a priori, is inferred from the unfolding of the dynamics during the reconstruction process. Indeed, the choice of lag,  $\lambda$ , and embedding dimension,  $d$ , determine whether the data is under-sampled (projected upon itself) or redundant; this is particularly important when dealing with a finite amount of noisy data. We note also that different observables of a system contain unequal amounts of dynamical information [9] implying that phase space reconstruction could be easier from one variable, but more difficult (or even impossible) from another. Our subsequent analysis is mindful of the balance between these caveats and the constraints imposed by the limited amount of noisy data.

Based on the PS reconstruction, various nonlinear measures have been defined to characterize process dynamics. As an illustration, we discuss three of these nonlinear measures, namely: (i) the first minimum in the mutual information function as a measure of decorrelation time, (ii) the correlation dimension as a measure of dynamic complexity, and (iii) the Kolmogorov entropy as a measure of predictability. We note that other nonlinear quantifiers, such as Lyapunov exponents, fractal dimensions, etc., have been also proposed and used [5, 6].

(i) The mutual information function (MIF) is a nonlinear version of the (linear) auto-correlation and cross-correlation functions, and was originally developed by Shannon and Weaver [10] with subsequent application to time series analysis by Fraser and Swinney [11]. The MIF measures the average information (in bits) that can be inferred from one measurement about a second measurement, and is a function of the time delay between the measurements. Univariate MIF measures predictability within the same data stream at different times. Bivariate MIF measures predictability of one data channel, based on measurements in a second signal at different times. For the present analysis, we use the first minimum in the univariate MIF,  $M_1$ , to indicate the average time lag that makes  $x_i$  independent of  $x_j$ . For a window of  $N$  points, we denote by  $Q$  the set of data measurements  $q_1, q_2, \dots, q_N$ , with associated occurrence probabilities  $P(q_1), P(q_2), \dots, P(q_N)$ . A second set,  $R$ , of data measurements,  $r_1, r_2, \dots, r_N$ , with a time delay relative to the measurements in the  $Q$  set, has occurrence probabilities  $P(r_1), P(r_2), \dots, P(r_N)$ . The function  $P(q_i, r_j)$  denotes the joint probability of both states occurring simultaneously. Then the MIF,  $I(Q, R)$ , the system entropy,  $H(Q)$  and the relative entropy,  $H(Q, R)$ , are defined by:

$$I(Q, R) = I(R, Q) = H(Q) + H(R) - H(Q, R), \quad (1)$$

$$H(Q) = -\sum_i P(q_i) \log[P(q_i)] \quad (2)$$

$$H(Q, R) = -\sum P(q_i, r_j) \log[P(q_i, r_j)]. \quad (3)$$

$H$  and  $I$  are expressed in units of bits if the logarithm is taken in base two.

(ii) The maximum-likelihood correlation dimension,  $D$ , is defined by [12, 13]:

$$D = \{(-1/M) \sum \ln[\delta_{ij} / \delta_0 - \delta_n / \delta_0] / (1 - \delta_n / \delta_0)\}^{-1}, \quad (4)$$

where  $M$  is the number of randomly sampled point pairs;  $\delta_{ij}$  is the maximum-norm distance between the (randomly chosen)  $i, j$  point pairs, as defined in Eq. 6 (below). The distance (scale length)  $\delta_n$  is associated with noise as measured from the time serial data. Note that the distances are normalized with respect to a nominal scale length  $\delta_0$  which is chosen as a balance between sensitivity to local dynamics (typically at  $\delta_0 \geq 5a$ ) and avoidance of excessive noise (typically at  $\delta_0 \geq a$ ). Here, the symbol  $a$  denotes the absolute average deviation as a robust indicator of variability [13] in the time serial data:

$$a = (1/N) \sum_{i=j}^N |x_i - \underline{x}|, \quad (5)$$

where  $\underline{x}$  is the mean of  $x_i$  over a window of  $N$  points. The distances  $\delta_{ij}$  are defined by:

$$\delta_{ij} = \max_{0 \leq k \leq m-1} |x_{i+k} - x_{j+k}|, \quad (6)$$

where  $m$  is the average number of points per cycle.

(iii) The Kolmogorov entropy,  $K$  measures the rate of information loss per unit time, or - equivalently - the degree of predictability. A positive, finite Kolmogorov entropy generally is considered to indicate that the underlying dynamics is chaotic. A very large entropy indicates a stochastic, non-deterministic (totally unpredictable) phenomenon. The entropy is calculated from the average time for two points on an attractor to pass from a small separation (lower than a certain threshold),  $\delta < \delta_0$ , to a separation larger than the threshold,  $\delta > \delta_0$ , [14]:

$$K = -f_s \log(1 - \underline{1/b}), = \underline{b}(1/M) \sum_{i=j}^M b_i, \quad (7)$$

where  $b_i$  is the number of time steps for two points, initially within  $\delta < \delta_0$ , to diverge to  $\delta > \delta_0$ . The symbol  $f_s$  denotes the data sampling rate, and  $M$  denotes the number of PS points in the average.

Entropy and correlation dimension usually are defined in the limit of zero scale length. However, all real data have noise and even noiseless model data are limited by the finite precision of computer arithmetic. Thus, we choose a finite scale length that is slightly larger than the noise ( $\delta_0 = 2a$ ), at which to report the values of  $K$  and  $D$ , corresponding to finite-scale dynamic structure. Consequently, the values of  $K$  and  $D$  that we report do not capture the full dynamical complexity and have smaller values than expected for the zero-scale-length limit ( $\delta_0 \rightarrow 0$ ).

### 3. DEFINITION OF THE PHASE SPACE DISSIMILARITY MEASURES

The traditional nonlinear measures (TNM) described in the previous section characterize global features of the nonlinear dynamics and distinguish sufficiently clearly between, say, regular and chaotic dynamics. However they do not saliently capture *slight* dissimilarities between dynamical states and therefore cannot distinguish between close dynamical regimes. The same is true for other global indicators, such as fractal dimension, Lyapunov exponents, etc. This lack of discrimination power occurs because the traditional measures are defined by averaging or integrating various quantities against the invariant distribution function over the attractor, thereby erasing many of the finer differences induced by the dynamics. Thus TNM can provide only one or a few scalar measures as summary descriptors of complex dynamics.

Greater discrimination is possible, in principle, by more detailed analysis of the reconstructed dynamics. In particular, the natural (or invariant) measure (distribution function) on the attractor contains more information and provides a more refined representation of the dynamics, insofar as it describes the visitation frequency of the various points on the attractor. Thus, instead of first integrating the invariant distribution function (thereby washing out the information it contains) and then assess the result, we first compare it with similar functions and *then* integrate. This approach - which is the crucial point of our method leads to significant enhancements of the discrimination power. To obtain a useful discrete representation of the invariant measure from time serial data, we proceed as follows. We first represent each signal value,  $x_i$ , as a symbolized form,  $s_i$ , that is one of  $S$  different integers,  $0, 1, \dots, S-1$ :

$$0 \leq s_i = INT[S(x_i - x_{min}) / (x_{max} - x_{min})] \leq S-1. \quad (8)$$

Here, the function  $INT$  converts a decimal number to the closest lower integer, and  $x_{min}$  and  $x_{max}$  denote the minimum and maximum values of  $x_i$ , respectively, over the base case (reference data). We previously used [15] the minimum and maximum values over both the basecase and testcase (data to be tested for departure from the basecase). However, in real- or near-real-time analyses, only basecase extrema are actually known while the testcase extrema are assumed to be equal to the former. We require that  $s_i(x_i = x_{max}) = S-1$  in order to maintain exactly  $S$  distinct symbols. Thus, the phase space is partitioned into  $S^d$  hypercubes or bins. By counting the number of phase-space points occurring in each bin, we obtain the distribution function as a discretized density on the attractor. We denote the population of the  $i$ -th bin of the distribution function,  $Q_i$ , for the base case, and  $R_i$  for a test case, respectively. The choice of parameters ( $S$ ,  $N$ ,  $d$ , and  $\lambda$ ) depends on the specific data under consideration as it reflects the features of the underlying dynamics.

We next compare the distribution functions of the testcase and basecase dynamics by measuring the difference between  $Q_i$  with  $R_i$  by the  $\chi^2$  statistics and  $L_1$  distance:

$$\chi^2 = \sum_i (Q_i - R_i)^2 / (Q_i + R_i), \quad (9)$$

$$L = \sum_i |Q_i - R_i|, \quad (10)$$

where the summations in both equations run over all of the populated cells in the PS. The choice of these measures is based on rather natural considerations. The  $\chi^2$  statistic is one of the most powerful, robust, and widely-used statistical tests to measure discrepancies between observed and expected frequencies [16]. The  $L_I$  distance is the natural metric for distribution functions since it is directly related to the total invariant measure on the attractor. These measures account for changes in the geometry of the attractor and for changes in the DF as well. To apply these measures properly, we scale the total population of the unknown distribution function (sum over all the domain populations in  $R_i$ ) to be the same as the total population of the basecase. The sum in the denominator of Eq. (10) is based on a test for equality of two multinomial distributions [16].

The previous analysis can be extended in a natural manner that is inherently compatible with the underlying dynamics. By connecting successive PS points as indicated by the dynamics,  $y(i) \rightarrow y(i+1)$ , one obtains a discrete representation of the process flow [4]. We thus form a  $2d$ -dimensional vector,  $Y(i)=[y(i), y(i+1)]$ , in the connected PS. As before,  $Q$  and  $R$  denote the distribution functions for the basecase and testcase, respectively, in the connected PS. We define the measure of dissimilarity between these two connected PS states, as before, via the  $L_I$ -distance and  $\chi^2$  statistics [15]:

$$\chi_c^2 = \sum_{ij} (Q_{ij} - R_{ij})^2 / (Q_{ij} + R_{ij}), \quad (11)$$

$$L_c = \sum_{ij} |Q_{ij} - R_{ij}|. \quad (12)$$

The subscript  $c$  indicates the connected distribution function measure. We note that the value  $\lambda=1$  results in  $d-1$  components of  $y(i+1)$  being redundant with those of  $y(i)$ , but we allow this redundancy to accommodate other data such as discrete points from two dimensional maps. Using pair-wise connectivity between successive  $d$ -dimensional states, this approach captures even more dynamical information. This additional information results in a higher discriminating power of the connected measures as compared with their non-connected counterparts. Indeed, the measures defined in Eqs. (10)-(13) are proven to satisfy the following inequalities [21]:

$$\chi^2 \leq L, \chi_c^2 \leq L_c, L \leq L_c, \chi^2 \leq \chi_c^2. \quad (13)$$

Henceforth we shall refer to the quantities defined in Eqs. (9 - 12) as phase space dissimilarity measures (PSDM), to differentiate them from the TNM, defined in Section 2.

#### 4. APPLICATION TO MODEL AND PHYSIOLOGICAL DATA

In the following examples, we compare the relative performance of two TNM, namely the correlation dimension,  $D$ , and Kolmogorov entropy,  $K$ , and two PSDM, namely the  $\chi^2$  statistic and the  $L_I$  distance. The disparate range, variability, and physical meaning of these indicators render direct comparison difficult. To achieve meaningful comparison, we convert the nonlinear measures into a consistent form, within a unified framework, by suitable renormalization [15, 18]. For each nonlinear indicator,  $V = D, K, \chi^2$ , and  $L$ , we define  $V_i$  as its value for the  $i$ -th cutset. We denote by  $\underline{V}$  the mean value of that indicator over the basecase, with a corresponding sample standard deviation  $\sigma$ . The renormalized form of the indicator is then  $U(V) = |V_i - \underline{V}| / \sigma$ , which measures the number of standard deviations that the testcase deviates from the basecase mean. It is expected that dynamical states close to the basecase have small values of the renormalized PSDM, while significant change is manifested by large values of the renormalized PSDM.

We assess the discriminating power of the new measures by several test examples. The first is the ordinary differential equation Lorenz model [19]:

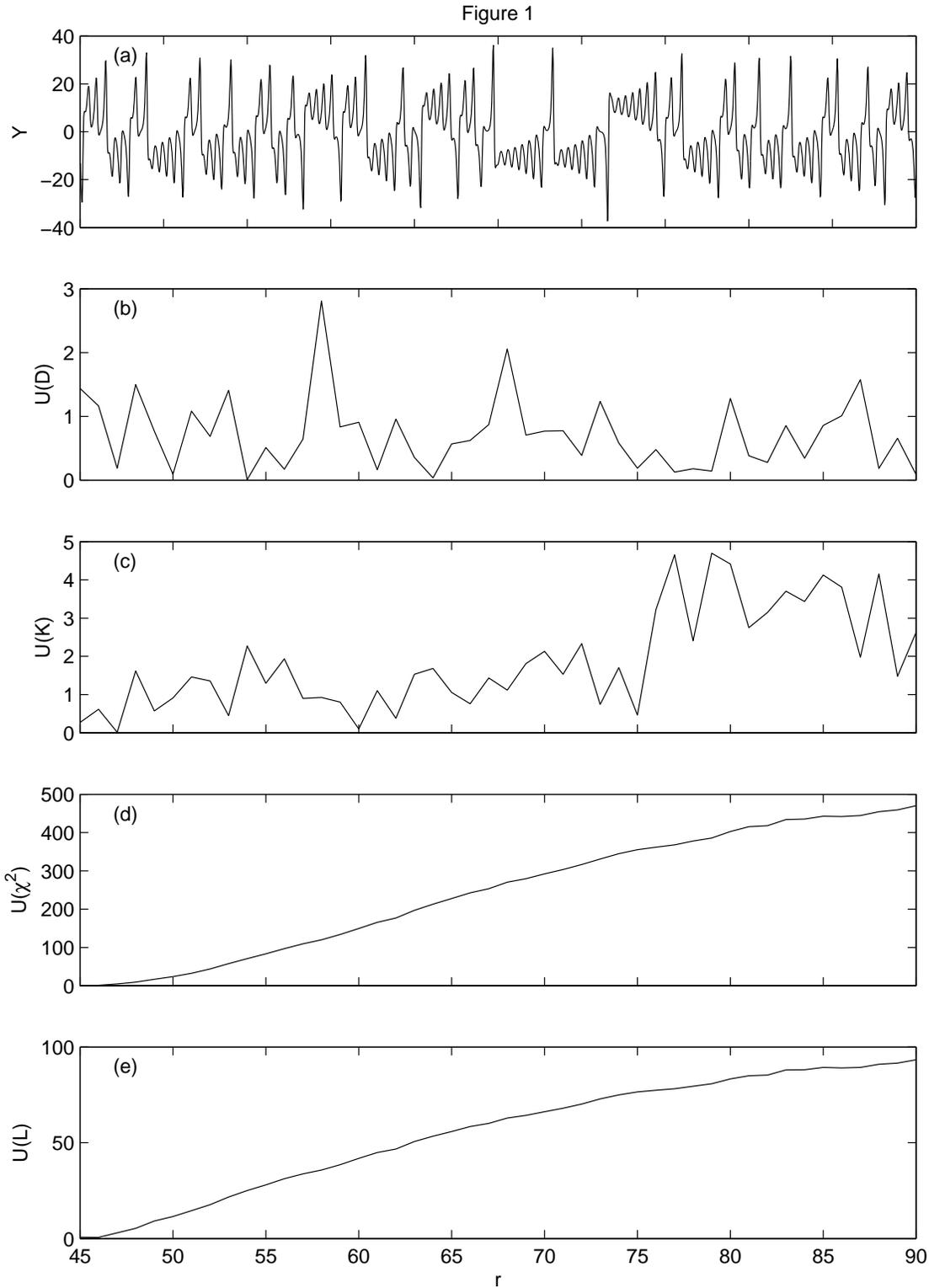
$$\dot{x} = \sigma(y - x), \quad \dot{y} = rx - y - xz, \quad \dot{z} = xy - bz. \quad (14)$$

The Lorenz model describes an oversimplified weather evolution in which  $\sigma = 10$ ,  $b = 8/3$  and  $r$  is a variable parameter. By varying  $r$ , one changes the type of stable asymptotic behavior of the Lorenz system from fixed points, to periodic orbits and eventually to a chaotic attractor. As mentioned before, TNM are good indicators of a bifurcation or transition to chaos. However, transitions between two chaotic regimes are not readily detected by these same measures, especially for relatively small changes in the parameter that underlies the transition. Therefore, we assessed the discriminating power of the PSDM by applying them to the systems above in region where they are known to behave chaotically, namely  $45 \leq r \leq 90$ . The results are displayed in Fig. 1.

The second example is the time-delayed ordinary differential equations “synthetic brain” model proposed by Bondarenko [20],

$$\dot{u}_i(t) = -u_i(t) + \sum_{j=1}^M a_{ij} f(u_j(t - \tau_j)), \quad i, j = 1, 2, \dots, M. \quad (15)$$

The Bondarenko model is a generalization of the Hopfield model for the electronic circuit realization of a neural network, which accounts for time delays,  $\tau_j$ . This model has been proposed and studied in connection to its capability to produce signals similar to human EEG signals. The state variable  $u_i(t)$  is the output signal of the  $i$ -th neuron and the matrix  $a_{ij}$  denotes the coupling coefficients between the neurons, with randomly chosen values,  $-2 \leq a_{ij} \leq 2$ . In our illustration, we use  $M = 10$  (ten neurons). The time delay of the  $j$ -th neuron output,  $\tau_j$ , is constant and equal to 10 conventional units. The nonlinear response function,  $f(x) = c \tanh(x)$ , simulates the nonlinear neural response to signals from neighboring neurons. By varying the coefficient  $c$  between 5 and 18, we change simultaneously the effective coupling between the neurons, while ensuring that



**Figure 1. Nonlinear measures versus  $r$  for the Lorenz system: (a) time serial data from channel  $y$ , (b) correlation dimension,  $D$ , (c) Kolmogorov entropy,  $K$ , (d)  $U(\chi^2)$ , and (e)  $U(L)$ . The phase space parameters are  $d=3$ ,  $S=20$ , and  $\lambda=2$ . Each cutest has  $N=50,000$  points.**

the system remains in chaotic regime. The results of this study are displayed in Fig. 2. The results in Figs 1 and 2 shows that the TNM indicate little change, if at all, while the PSMD unambiguously track and robustly display the underlying condition change.

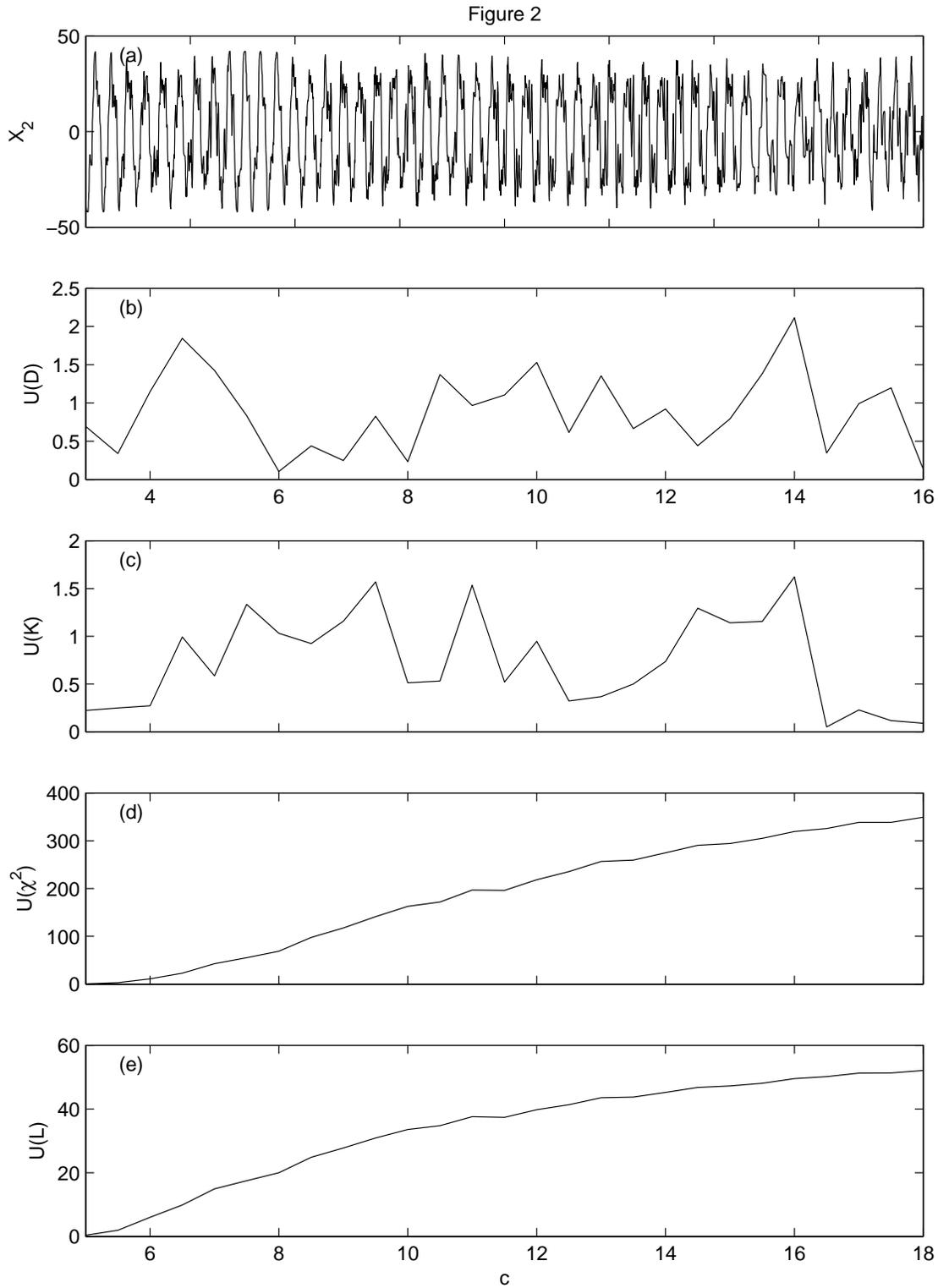
Having assessed the performance of the new measures on model data, we applied them to four sets of different clinical data, as briefly described below.

Human electroencephalogram (EEG) data were acquired during clinical epilepsy monitoring and analyzed according to the procedure described in Sections 2 and 3. Figure 3 shows typical results. Raw EEG in subplot (a) has very complex, non-periodic features that are typical of brain waves. The seizure event occurred at 110.7 minutes, as denoted by the solid vertical line in subplots (d) and (e). No seizure event forewarning is provided by the correlation dimension in subplot (b), or by the Kolmogorov entropy in subplot (c). The isolated peaks at 42 minutes and 58 minutes in subplot (c) are not significant. More than 27 minutes of seizure-event forewarning are provided by both  $U(\chi^2)$  in subplot (d) and  $U(L)$  in subplot (e), which have two (or more) successive occurrences above the threshold of 5 (the dashed horizontal line) at 85 minutes (the vertical dashed line). We have demonstrated this approach for seizure forewarning in 10-61 datasets [15, 21, 18].

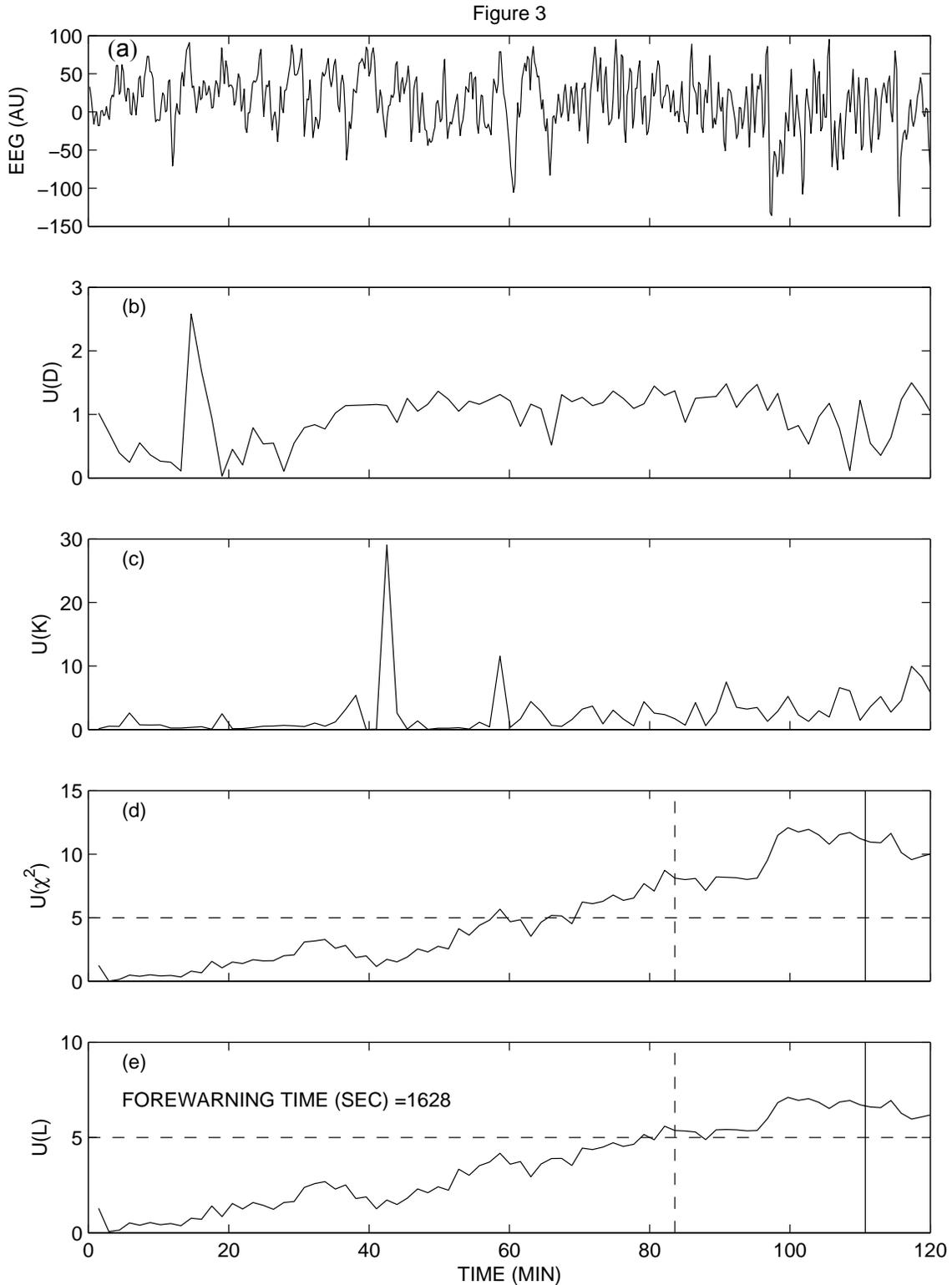
Human electrocardiogram (ECG) data were acquired during ambulatory monitoring; see Ref. [22] for details. Figure 4 shows results for a ventricular fibrillation event at 37 minutes. Raw ECG in subplot (a) has ten distinct heartbeats and their associated quasi-periodic (nonlinear) features. The correlation dimension in subplot (b) varies randomly without any forewarning features, showing a rise at the fibrillation event. The Kolmogorov entropy in subplot (c) varies erratically. The peaks occurring at 8 and 24 are not valid forewarning indications. Sixteen minutes of fibrillation-event forewarning (the vertical dashed line) are provided by both  $U(\chi^2)$  in subplot (d) and  $U(L)$  in subplot (e), which have two (or more) successive occurrences above the threshold (the dashed horizontal line).

A surface stethoscope acquired lung sounds data during lung experiments on anesthetized pigs [23]. The nominal state consisted of normal breathing. Subsequent test cases were obtained by injecting a controlled volume of air (in increments of 100 milliliters up to 1400 milliliters) in the space between the diaphragm and the lungs, making breathing increasingly more difficult. Figure 5 shows sample pneumothorax results. Raw lung sounds data in subplot (a) has very complex features, including quasi-periodic heartbeats that are superimposed on breath-cycle undulations. The correlation dimension in subplot (b) provides no clear indication of condition change. The Kolmogorov entropy in subplot (c) likewise varies erratically. Condition change is indicated by both  $U(\chi^2)$  in subplot (d) and  $U(L)$  in subplot (e), which rise to a plateau of 5 over 100-500 ml, then increase to values larger than 20 over 500-1300 ml thereby providing robust forewarning of the animal's death, which occurred at 1400 ml.

Finally, electrocardiogram (ECG) data were obtained from anesthetized rats subjected to an induced sepsis experiment. After 55 minutes of normal-state recording, the rat was exposed to inhaled bacterial endotoxin that induces an inflammatory response and eventually sepsis. Figure 6 shows sample results. Raw ECG in subplot (a) has 14 distinct heartbeats with additional quasi-periodic (nonlinear) features. No indication of condition change is displayed by either the

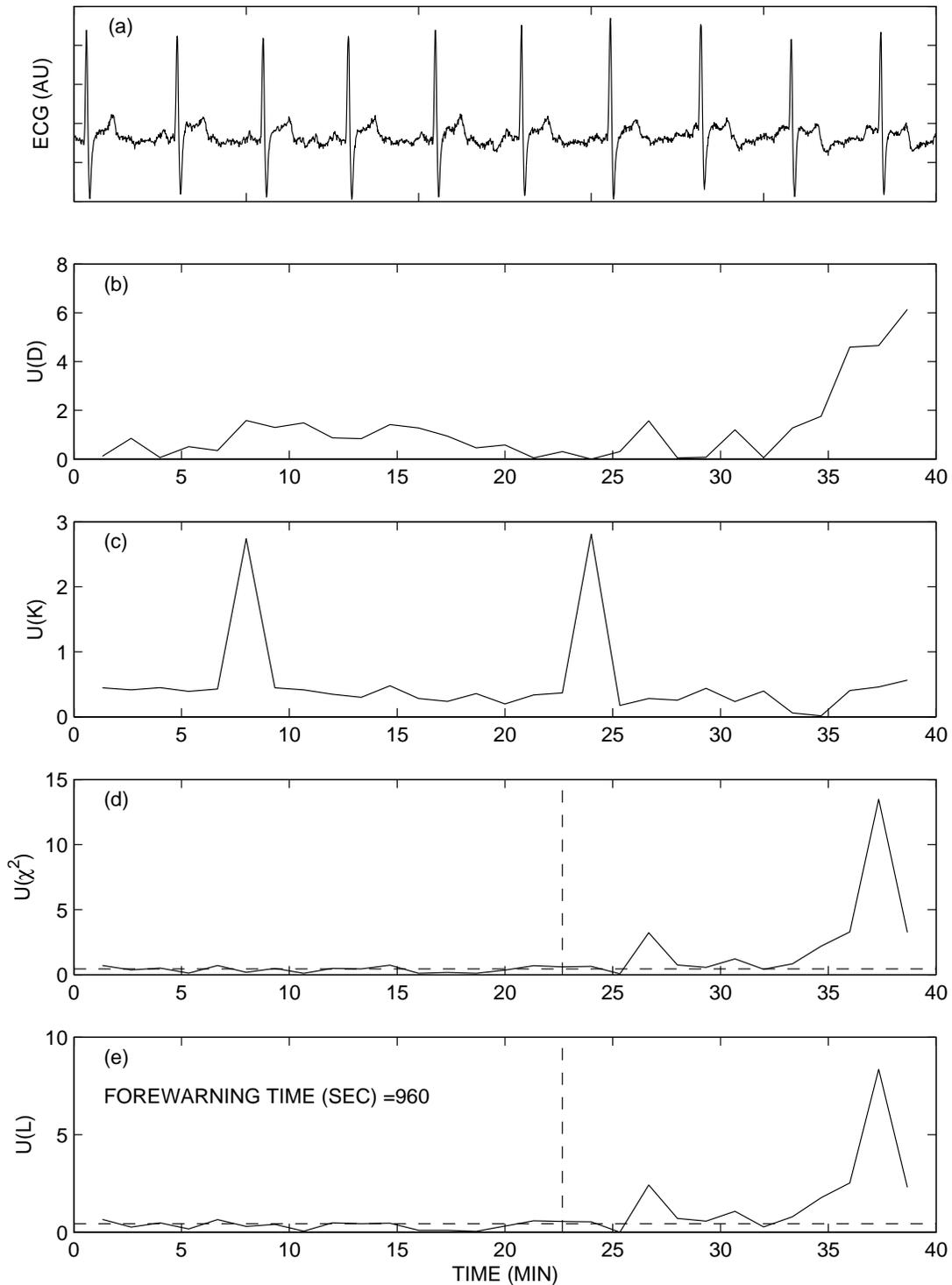


**Figure 2. Nonlinear measures versus  $c$  for one neuron channel in the Bondarenko system: (a) time serial data from neuron 2, (b) correlation dimension,  $D$ , (c) Kolmogorov entropy,  $K$ , (d)  $U(\chi^2)$ , and (e)  $U(L)$ . The phase space reconstruction parameters are  $d = 3$ ,  $S = 7$  and  $\lambda = 1$ . Each cutest has  $N = 20,000$  points.**

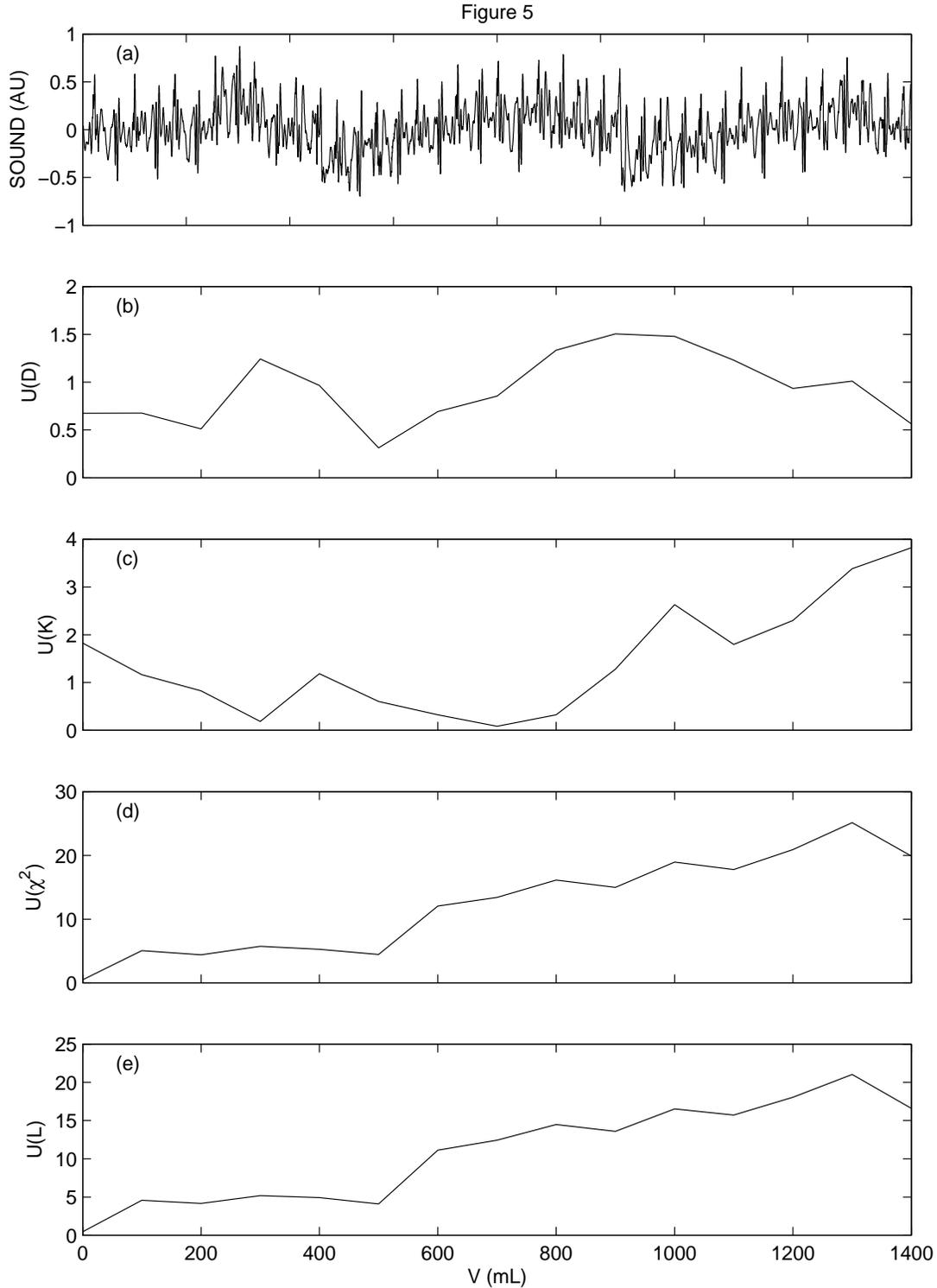


**Figure 3. Results for Channel 5 of dataset #PVM006, showing time-series plots for: (a) 2.4 seconds of raw EEG data collected at 250Hz, (b) correlation dimension,  $D$ , (c) Kolmogorov entropy,  $K$ , (d)  $U(\chi^2)$ , and (e)  $U(L)$ . The phase space dissimilarity measures in subplots (d) and (e) were computed for  $d = 3$ ,  $S = 20$ ,  $\lambda = 17$ , and after removal of eye blink artifacts. Each cutest has  $N = 22,000$  points, corresponding to 88 seconds.**

Figure 4

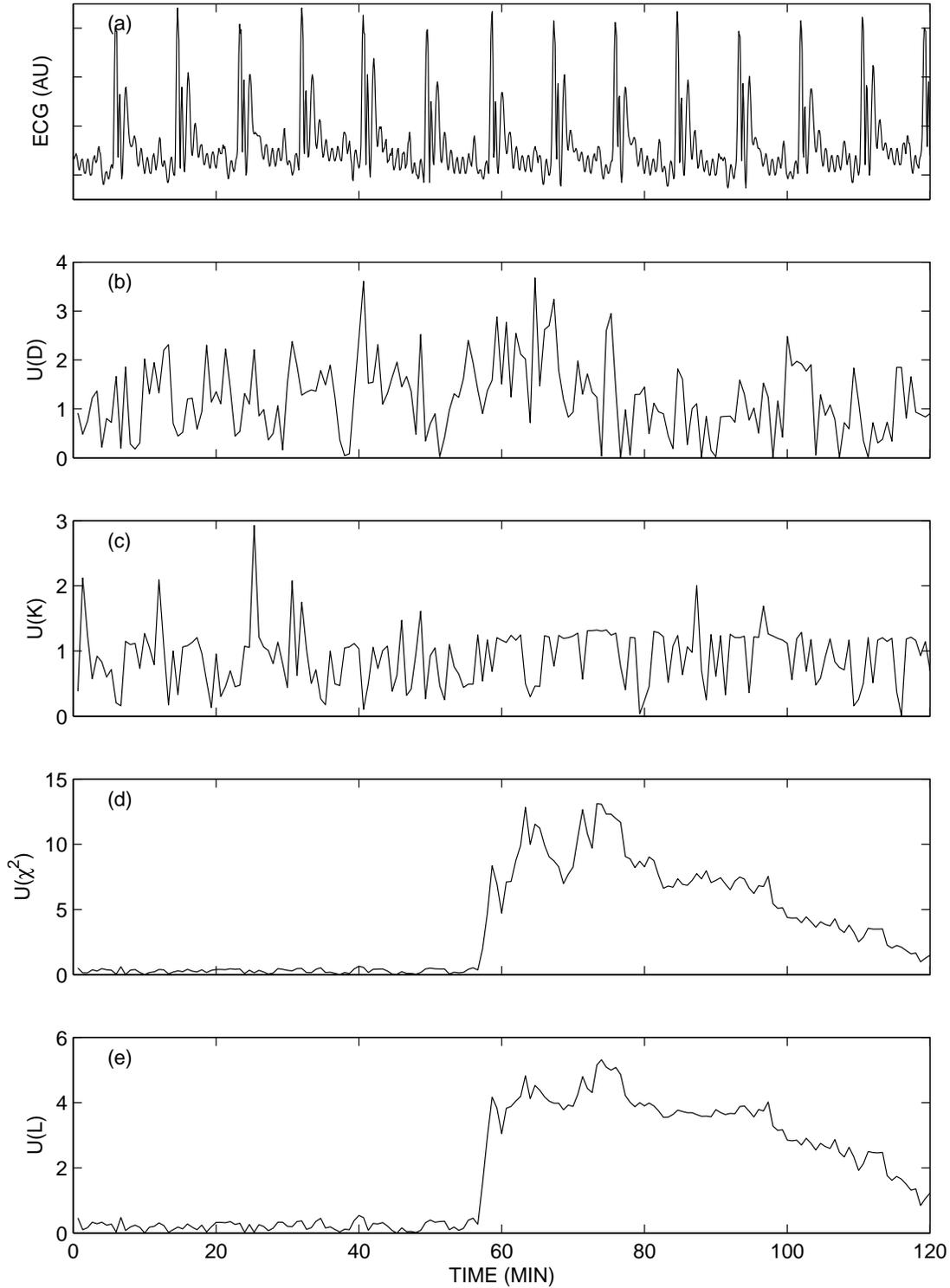


**Figure 4. Results for dataset #EC8202, showing time-serial plots for: (a) 10 seconds of raw ECG data collected at 250 Hz, (b) correlation dimension,  $D$ , (c) Kolmogorov entropy,  $K$ , (d)  $U(\chi^2)$ , and (e)  $U(L)$ . The phase-space dissimilarity measures in subplots (d) and (e) were computed for  $d=5$ ,  $S=3$ ,  $\lambda=27$ , after removal of breathing artifacts. Each cutest had  $N=18,000$  points, corresponding to 72 seconds.**



**Figure 5. Results from dataset #PTX5, showing time-serial plots for: (a) 4 seconds of raw lung sounds data collected at 10 kHz, (b) correlation dimension,  $D$ , (c) Kolmogorov entropy,  $K$ , (d)  $U(\chi^2)$ , and (e)  $U(L)$ . The phase-space dissimilarity measures in subplots (d) and (e) were computed for  $d = 3$ ,  $S = 30$ ,  $\lambda = 20$ , after removal of breathing artifacts. Each cutest has  $N = 50,000$  points, corresponding to 5 seconds.**

Figure 6



**Figure 6. Results for dataset #33209V, showing time-series plots for: (a) 2.4 seconds of ECG data collected at 500 Hz, (b) correlation dimension,  $D$ , (c) Kolmogorov entropy  $K$ , (d)  $U(\chi^2)$ , and (e)  $U(L)$ . The phase-space dissimilarity measures in subplots (d) and (e) were computed for  $d=2$ ,  $S=2$ ,  $\lambda=80$ , after removal of breathing artifacts. Each cutest has  $N=20,00$  points, corresponding to 40 seconds.**

correlation dimension in subplot (b), or by the Kolmogorov entropy in subplot(c). The condition change is shown clearly by both  $U(\chi^2)$  in subplot (d) and  $U(L)$  in subplot (e), which remain low for the first 57 minutes, rising abruptly after the exposure onset, remaining high for the next 20 minutes, then decreasing slowly as the immune response fought off the bioagent effects. This recovery response is consistent with other physiological observations during the test (not shown).

## 5. DISCUSSION

We have developed model-independent indicators to detect condition change (dissimilarity) in nonlinear time series. The phase space indicators of condition change measure the difference between suitably discretized forms of the invariant distribution functions on the attractor for the basecase and testcase, as  $\chi^2$  statistics and  $L_1$  distance. Using these metrics magnifies the differences between the process dynamics, and avoids the inner cancellation effects due to averaging which occurs when calculating the TNM. As a result, significant changes in model dynamics are clearly detected by the PSDM, as the parameters vary. On the other hand, these changes are either barely detected or undetected by the TNM, such as the Kolmogorov entropy or the correlation dimension. For real physiological data, the same conclusion can be drawn, although the difference is less marked, due to inherent noise. Indeed, the two model examples show a gradual rise in PSDM as the underlying parameter increases monotonically with time, as shown in Figs. 1-2. The four physiological examples show small values of the PSDM in the normal state, followed by a notable monotonic rise and/or by successive occurrences of the PSDM above the “normality threshold” that indicates forewarning or onset of abnormal (pathological) dynamics, as shown in Figs. 3-6. In all applications, the renormalized PSDM show results that are consistently more robust and more timely than those provided by the TNM.

An early version [17] of this approach was successfully applied to detect dynamical change in various physical processes. Examples include: distinguishing different drilling conditions from spindle motor current of a machining center; detecting balanced and unbalanced centrifugal pump conditions from motor power; and predicting failure of a bellows coupling in a rotating drive train from motor current. More recent, improved analyses include: discerning the difference in microcantilever vibrations with and without mercury on the sensor and forewarning of various machine and equipment failure [24]. Success for such diverse applications suggests that this technique can be reliably used for measuring condition change in nonlinear processes.

## 6. ACKNOWLEDGMENTS

We gratefully acknowledge support of the basic forewarning methodology research by the U.S. Department of Energy's Office of Basic Energy Sciences and by the U.S. Department of Energy's Nuclear Energy Research Initiative (NERI2000-109). We thank Nicolet Biomedical Inc. (now ViaSys Healthcare Inc.) of Madison, Wisconsin, for sponsorship of epilepsy forewarning research under a Cooperative Research and Development Agreement (CRADA # ORNL 99-0559). Biomedical research also has been sponsored by the Laboratory Directed Research and

Development Program of Oak Ridge Nation Laboratory, managed by UT-Battelle, LLC, for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725, for forewarning of epileptic events, condition change from lung sounds, and sepsis forewarning. Research on cardiac event forewarning has been sponsored by Physio-Control Corp. under a Cooperative Research and Development Agreement (CRADA # ORNL 95-0353).

## REFERENCES

- [1] J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57**, 617 (1985).
- [2] E. Atlee Jackson, *Perspectives of Nonlinear Dynamics*, Cambridge University Press, Cambridge, vol. 1989; vol. 2, 1990.
- [3] S. H. Strogatz, *Nonlinear Dynamics and Chaos*, Addison-Wesley, Reading, 1994.
- [4] H. D. I. Abarbanel, *Analysis of Observed Chaotic Data*, Springer Publ., New York, 1996.
- [5] *Proceedings of the Workshop on "Chaos in Brain ? "*, March 10-12, 1999, K. Lehnertz, J. Arnhold, P. Grassberger, and C. E. Elger, eds., World Scientific, Singapore, 2000.
- [6] D. T. Kaplan and L. Glass, *Understanding Nonlinear Dynamics*, Springer, New York, 1995; Focus Issue: Mapping and Control of Complex Cardiac Arrhythmias, *Chaos*, **12**, no. 3 (2002).
- [7] L. M. Hively, N. E. Clapp, C. S. Daw, and W. F. Lawkins, "Nonlinear Analysis of EEG for Epileptic Seizures," ORNL/TM-12961, 1995.
- [8] F. Takens, *Lecture Notes in Mathematics*, **898**, 366 (1981) Springer, New York.
- [9] C. Letellier, J. Maquet, L. Le Sceller, G. Gouesbet, and L. A. Aguirre, *J. Phys. A*, **31**, 7913 (1998).
- [10] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, 1949.
- [11] A. M. Fraser and H. L. Swinney, *Phys. Rev A*, **33**, 1134-1140 (1986).
- [12] F. Takens, *Lecture Notes in Mathematics*, **1125**, 99-106 (1984) Springer-Verlag, Berlin.
- [13] J. C. Schouten, F. Takens, and C. M. van den Bleek, *Phys. Rev. E*, **50**, 1851 (1994).
- [14] J. C. Schouten, F. Takens, and C. M. van den Bleek, *Phys. Rev. E*, **49**, 126 (1994).

- [15] L. M. Hively, P. C. Gailey, and V. A. Protopopescu, *Phys. Lett. A*, **258**, 103 (1999); L. M. Hively, P. C. Gailey, and V. A. Protopopescu, in *Proc. Int. Workshop: "Chaos in Brain ?"*, Bonn, Germany, 10-12 March, 1999, K. Lehnertz et al. eds., World Scientific, 2000, p. 333; and P. C. Gailey, L. M. Hively, V. A. Protopopescu, in *Proc. Fifth Exp. Chaos Conf. (ECC5)*, Orlando, Florida, June 28 - July 1, 1999, to appear.
- [16] A. M. Mood, S. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, McGraw Hill (1974).
- [17] L. M. Hively, in *Proc. Maintenance and Reliability Conf.*, edited by T.E. Shannon et al. (Univ. of Tennessee, Knoxville), **1**, 16.01 (1997).
- [18] V. Protopopescu, L. M. Hively, and P. C. Gailey *J. Clin. Neurophys*, **18**, 223 (2001).
- [19] E. N. Lorenz, *J. Atmos. Sci*, **20**, 130 (1963).
- [20] V. E. Bondarenko, *Int. J. Bifurc. Chaos*, **7**, 1133 (1997).
- [21] L. M. Hively, V. Protopopescu, and P. Gailey, *Chaos*, **10**, 864 (2000).
- [22] N. E. Clapp, L. M. Hively, and R. E. Stickney, "CRADA Final Report for CRADA Number ORNL 95-0353 - Heart Pathology Determination from Electrocardiogram Signals by Application of Deterministic Chaos Mathematics," ORNL/TM-13436, 1996.
- [23] R. C. Ward, K. L. Kruse, G. O. Allgood, L. M. Hively, K. N. Fischer, N. B. Munro, and C. E. Easterly, "Virtual Human Project" in "Visualization of Temporal and Spatial Data for Civilian and Defense Applications," G.O. Allgood and N. L. Faust, eds. *Proc. SPIE* **4368**, 158 (2001).
- [24] L. M. Hively and V. Protopopescu, "Forewarning of Failure in Critical Equipment at Next Generation Nuclear Power Plants," ORNL/TM-183, 2002.