ROBUST DETECTION OF DYNAMICAL CHANGE IN SCALP EEG

PAUL C. GAILEY, LEE M. HIVELY, and VLADIMIR A. PROTOPOPESCU
Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

ABSTRACT

We present a robust, model-independent technique for measuring changes in the dynamics underlying nonlinear time-serial data. We define indicators of dynamical change by comparing distribution functions on the attractor via L_1-distance and χ^2 statistics. We apply the measures to scalp EEG data with the objective of capturing the transition between non-seizure and epileptic brain activity in a timely, accurate, and non-invasive manner. We find a clear superiority of the new metrics in comparison to traditional nonlinear measures as discriminators of dynamical change.

1. Introduction

This work focuses on nonlinear analysis of physiological data. Typically, these data arise from a virtual "black box" with little knowledge of the underlying system, its dimensionality, or noise contamination. More often than not, nonlinear analysis requires some assumptions about the underlying dynamics. For example, calculation of Lyapunov exponents or Kolmogorov entropy implicitly assumes that the physiological process can be modeled as a dynamical system. At a more fundamental level, one may ask whether the data arises from a stationary process. Numerous statistical tests for stationarity exist, but such tests usually assume that the dynamics are stationary within the two time windows under comparison. Moreover, complex systems, like the brain, may not be well modeled by stationary dynamics over long times.

We describe a model-independent method for measuring change in nonstationary data. The dynamics of reference and test cases are represented as discrete distributions of the density of points in reconstructed phase space during different time windows. Variability is captured by the visitation frequency at various regions of phase space as described by the distribution function. The method quantifies differences in these reconstructed dynamics by comparing the distribution functions. We make no assumption about stationarity, because no dynamical properties are inferred from the reconstructed attractor. The system dynamics may change within the time window, but such variability presents no problem for our technique, which measures dynamical change over a variety of length scales, and over a wide range of time scales. Moreover, our method allows measurement of dynamical change that that occurs continuously or intermittently.

Recently, Moeckel and Murray[1] discussed similar concepts for measuring the "distance" between attractors from time-delay reconstructions. In this context, our method provides continuous measures of change in contrast to stationarity tests for whether or not *any* statistically significant change has occurred. Due to their continuous nature and their independence from assumptions about stationarity, our measures are particularly useful for analysis of physiological data.

2. Traditional Nonlinear Measures

We assume that an unknown scalar signal, x, is sampled at equal time intervals, τ, starting at time, t_0, yielding a sequence of N points, $x_i = x(t_0 + i\tau)$. Dynamical process reconstruction[2] uses d-dimensional time-delay vectors, $y(i)=[x_i, x_{i+\lambda}, \ldots, x_{i+(d-1)\lambda}]$, for a system with d active variables and time lag, λ. The choice of lag and embedding dimension, d, determine how well the reconstruction unfolds the dynamics for a finite amount of noisy data. A proper reconstruction allows calculation of nonlinear measures that are consistent with the original dynamics. Below, we use three traditional measures, for comparison to our phase-space indicators of dissimilarity.

The mutual information function is a nonlinear form of auto-correlation function. Mutual information was devised by Shannon and Weaver[3], and applied to time series by Fraser and Swinney[4]. Mutual information measures the information (in bits) that can be inferred from one signal about a second signal, and is a function of the time delay between the measurements. Univariate (bivariate) mutual information measures information within the same (different) data stream(s) at different times. Here, we use the first minimum, M_1, in the univariate mutual information function. M_1 measures the average time separation (in timesteps) that decorrelates two points in the time series.

The correlation dimension measures process complexity and is a function of scale length, δ, in the data. Our choice of length scale balances local dynamics (typically at $\delta \leq 3a$) against avoidance of excessive noise (typically at $\delta \geq a$). The symbol, a, denotes the absolute average deviation as a robust indicator of variability[5] in the time serial data:

$$a = (1/N) \sum_{i=1}^{N} |x_i - \underline{x}|. \qquad (1)$$

The symbol $\underline{x}$ denotes the mean of x_i. We use the maximum-likelihood correlation dimension, D, developed by Takens[6] with modifications for noise by Schouten et al.[5]

The Kolmogorov entropy, K, measures the rate of information loss (bits/s). Positive, finite entropy generally is considered to clearly indicate chaotic features. Large entropy implies a stochastic, totally unpredictable process. Entropy measures the average time for two points on an attractor to evolve from a small initial separation to more than a specific (large) distance, $\delta > \delta_0$. We use maximum-likelihood entropy by Schouten et al[7].

Noise corrupts all real data. Also, finite precision computer arithmetic truncates model data. Thus, we choose a finite-scale length that is larger than the noise, $\delta_0 = 2a$, at which to report K and D, corresponding to finite-scale dynamical structure. Thus, our values of K and D have smaller values than expected for the zero-scale-length limit.

3. New Measures of Dynamical Change

Traditional nonlinear measures characterize global features by averaging or integrating over the data. Such measures describe the long-term behavior but poorly indicate dynamical change. Greater discrimination is possible by more detailed analysis of the reconstructed dynamics. The natural (or invariant) measure on the attractor

provides a more refined representation of the reconstruction, describing the visitation frequency of the system dynamics over the phase space.

We converted each signal value, x_i, to one of S different integers, {0, 1, …, S-1}:

$$0 \le s_i = \mathrm{INT}[S(x_i - x_{min})/(x_{max} - x_{min})] \le S-1. \quad (2)$$

Here, x_{min} and x_{max} denote the minimum and maximum values of x_i, respectively, over both the reference case and over the test cases. INT is a function that converts a decimal number to the next lower integer. For $x_{min} \le x_i \le x_{max}$, the inequality $0 \le s_i \le S-1$ holds trivially. We took $s_i(x_i = x_{max}) = S-1$ in order to maintain exactly S distinct symbols and to partition the phase space into S^d hypercubes or bins. We then discretized the distribution function on the attractor, by counting the number of phase-space points occurring in each bin. We denoted the population of the i-th bin of the distribution function, P_i, for the base case, and Q_i for a test case, respectively. For this initial work, we iteratively varied each parameter (S, d, N, etc.) with the others fixed, to obtain optimum sensitivity of the measures to changes in EEG dynamics. A systematic method to determine optimal values for these parameters is the subject of future work.

We used an embedding window, $M_1 = (d - 1)\lambda$. Here, the first minimum in the mutual information function, M_1, is measured in timesteps. We obtained an integer value for the reconstruction lag by setting $\lambda = \mathrm{INT}[0.5 + M_1/(d-1)] \ge 1$, thus constraining the largest value of dimensionality to $d \le 2M_1 + 1$.

We compared the distribution function of a test state to the reference state, by measuring the difference between P_i with Q_i via the χ^2 statistics and L_1 distance:

$$\chi^2 = \sum_i (P_i - Q_i)^2/(P_i + Q_i), \text{ and} \quad (3)$$

$$L = \sum_i |P_i - Q_i|. \quad (4)$$

The summations include all of the populated cells in the phase space. The sum in the denominator of Eq. 3 is based on a test for equality of two multinomial distributions. Proper application of these measures requires a rescaling so that the total population of the test case distribution function is the same as the total population of the base case.

By connecting successive phase-space points as indicated by the dynamics, $y(i) \rightarrow y(i+1)$, we constructed a 2d-dimensional phase-space vector, $Y(i)=[y(i), y(i+1)]$. Thus, we obtained a discrete representation of the process flow.[8] This approach extends the method to capture more dynamical information using pair-wise connectivity between successive d-dimensional states.

We use base S arithmetic to assign an identifier $j \equiv I_i$ for the i-th phase-space state, using $I_i = \Sigma d^{m-1} s_i(m)$. The sum runs from m=1 to m=d, corresponding to successive components of the d-dimensional phase-space vector. The symbol, $s_i(m)$, denotes the m-th component of the i-th phase-space vector. The numeric identifier for the sequel phase-space point is $k \equiv I_{i+1}$. Then, we can define the measure of the dissimilarity between these two connected phase-space states, as before, via the L_1-distance and χ^2 statistics:

$$\chi_c^2 = \sum_{jk} (P_{jk} - Q_{jk})^2/(P_{jk} + Q_{jk}), \text{ and} \tag{5}$$

$$L_c = \sum_{jk} |P_{jk} - Q_{jk}|. \tag{6}$$

P_{jk} and Q_{jk} denote the distribution functions for the basecase and testcase, respectively, in the connected phase space. The summations in both equations run over all of the populated cells in the connected phase space. The subscript, c, denotes the connected measures, which are stronger metrics than the non-connected versions, according to the following inequalities[9-10]: $\chi^2 \leq L$, $\chi_c^2 \leq L_c$, $L \leq L_c$, and $\chi^2 \leq \chi_c^2$.

We tested the discriminating power of our measures on chaotic regimes of the Lorenz system[9] and of the Bondarenko model.[10-11] The latter model mimics high-dimensional EEG dynamics via a system of delay-differential equations. Over a broad parameter range, the phase-space measures increased monotonically by more than four orders of magnitude. Over this same range, traditional nonlinear measures were indistinguishable from noise or varied erratically by a factor of two. These results gave us confidence that the phase-space measures would be useful for noisy clinical EEG data.

4. EEG Analysis and Results

We converted one channel of analog EEG data on VHS tapes to 12-bit digital form at a sampling rate of 512 Hz. We chose N=20480 data points for each cutset. This choice balances better time discrimination (smaller N) against higher statistical power (larger N). We used the first 400 seconds of data to construct ten non-overlapping 40-second basecase cutsets. We compared each base case cutset to every test case cutset to obtain average values for χ^2 and L (and a corresponding standard deviation of the mean). We overlapped adjacent test case cutsets by 50% for smooth time-history trending. We also removed muscular artifacts (e.g., eye blinks) with a zero-phase quadratic filter.[9-10]

We found that d=3 and S=34 were adequate for our EEG data. The value of M_1 came from the first 400 seconds of (nonseizure) data. However, the disparate range and variability of the conventional and phase-space measures were difficult to interpret. Thus, we renormalized the nonlinear measures. For each nonlinear measure, V, we defined V_i as the value of nonlinear measure for the i-th cutset. The variable, V, was in turn D, K, M_1, χ^2, etc. We obtained the mean, $\underline{V}$, of V_i over ten non-overlapping cutsets (each with N=20480) for the first 400 seconds (base case interval) of the dataset. The corresponding sample standard deviation was denoted by σ. Then, the renormalized form was $U(V) = |V_i - \underline{V}|/\sigma$. For an indication of change, we used $U > U_c = 4.265$, corresponding to a false positive probability of $<10^{-5}$ in Gaussian random data. We required two or more consecutive occurrences of a positive indication to avoid spurious false positives, corresponding to a joint false positive probability of $<10^{-10}$ in Gaussian data. We next illustrate the use of these renormalized forms for measuring changes in scalp EEG.

Figure 1 shows results for one dataset. The first 300 seconds of data displays modest variability in all of the measures, representing the dynamics of normal brain activity. The clinical seizure occurs at 966-1035 seconds, as indicated by the vertical bars

at these times; all of the measures clearly show the seizure. Maxima and minima in the raw EEG (Fig. 1a) provide no preseizure indications, nor does the correlation dimension (Fig. 1b). Both Kolmogorov entropy (Fig. 1c) and mutual information (Fig. 1d) show preseizure change, beginning at 750 seconds. Connected phase-space measures (Figs. 1e-1f) have significant dissimilarity, beginning at 600 seconds. Measures of dissimilarity for the non-connected phase-space become really large, beginning at 700 seconds.

Table 1 summarizes the forewarning times (T in seconds prior to the clinical seizure) for each measure over nine EEG datasets. A negative value of T corresponds to an indication after onset of the seizure. Analysis of normal EEG shows no positive indication of change. M_1 provides the largest average forewarning, but gives no pre-seizure indication in three datasets. Correlation dimension gives the least pre-seizure warning, with no preseizure indication in one dataset. Kolmogorov entropy gives an intermediate level of forewarning with no preseizure indication in one instance. We denote the average time of preseizure indications as $\underline{T}$ for each nonlinear measure. The connected phase-space measures give more forewarning than the non-connected ones: $\underline{T}(L_c) - \underline{T}(L) = 144$ seconds, and $\underline{T}(\chi_c^2) - \underline{T}(\chi^2) = 131$ seconds, with indications for all datasets. We conclude that the phase-space measures are superior to the conventional nonlinear measures as indicators of dissimilarity.

Table 1: Times (in seconds prior to seizure) at which condition change is detected

Dataset #	D	K	M_1	L_c	L	χ_c^2	χ^2	best
109310	**1155**	1135	*	**1155**	-45	**1115**	-45	1115
109314	1110	1940	1700	1880	1880	**1960**	**1960**	1960
119230	371	491	451	**911**	**911**	**911**	**911**	911
119234	2060	2060	**2120**	**2120**	**2120**	**2120**	**2120**	2120
62723t	-100	1380	1400	**1720**	**1720**	**1720**	**1720**	1720
69212	716	**1496**	996	736	736	736	736	1496
73305D	-55	-55	*	**785**	**785**	**785**	**785**	785
c8492D	-14	206	166	**326**	226	**326**	306	326
wm12sD	*	*	*	**521**	**521**	**521**	**521**	521
Maximum time	2060	2060	2120	2120	2120	2120	2120	
Minimum time	-100	-55	166	326	226	326	306	
Average time	655	1082	1139	1128	984	1133	1002	
# false negatives	1	1	3	0	0	0	0	

Entries denoted by an asterisk (*) show no positive indications of condition change.
For each dataset, bold entries denote the earliest time of dissimilarity indication.

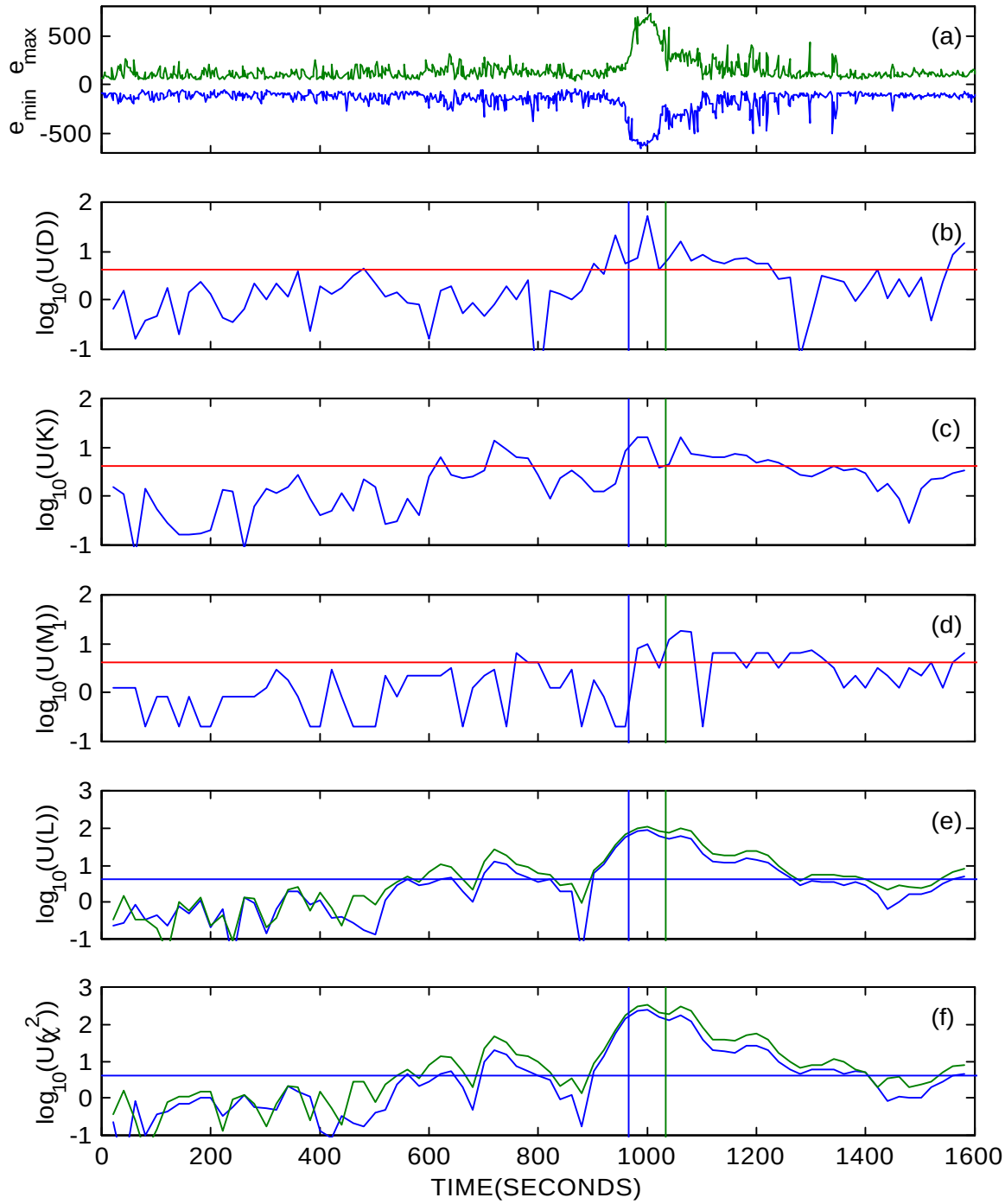

Figure 1: Results for dataset #c8492D versus time: (a) maxima (e_{max}) and minima (e_{min}) in the raw EEG, (b) change metric for correlation dimension, U(D), (c) change metric for Kolmogorov entropy, U(K), (d) change metric for first minimum in mutual information function, $U(M_1)$, (e) change metric for connected (solid) and non-connected (- -) L_1, and (f) change metric for connected (solid) and non-connected (- -) χ^2 . The ordinate values of the change metric (U) are in units of standard deviations from the mean.

5. Conclusions

These results have several new and important features. First, we use *non-invasive scalp EEG* in contrast to previous studies[12-13] that used data from subdural and depth electrodes. Second, earlier effort obtained EEG from the epileptogenic area. We use only *channel 13* data in the bipolar montage (closest to the patients' right eye), showing the robustness of our method. Third, prior effort used invasive monitoring to avoid low-frequency artifacts, which we remove from scalp data with a novel *zero-phase quadratic filter*. Fourth, previous investigations used conventional nonlinear measures, which we earlier found to provide no consistent preseizure trends.[14] Instead, we focus on model-independent *phase-space measures* of dissimilarity. By their definition, these indicators better capture the *differences* between the process dynamics, avoiding cancellation effects due to averaging over many orbits (e.g., correlation dimension and Kolmogorov entropy). Our measures provide preseizure indications in all nine EEG datasets. In sharp contrast, the traditional nonlinear measures provide inferior or no preseizure indications. Useful epilepsy forewarning will require analysis of several seizures for each patient, and detailed determination of detection criteria. Future work will evaluate: (i) false positives and negatives in epileptic patients, (ii) false positives in normal patients, and (iii) optimal reconstruction parameters, and (iv) the earliest indication of condition change.

6. References

1. B. R. Moeckel and B. Murray, *Physica D* **102** (1997) 187.
2. J.-P. Eckmann and D. Ruelle, *Rev. Mod. Phys.* **57** (1985) 617.
3. C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, (University of Illinois Press, Urbana, 1949).
4. A. M. Fraser and H. L. Swinney, *Phys. Rev A* **33** (1986**)** 1134.
5. J. C. Schouten, F. Takens, and C. M. van den Bleek, *Phys. Rev. E* **50** (1994) 1851.
6. F. Takens, *Lecture Notes in Mathematics* **1125** (1984) 99 (Springer-Verlag, Berlin).
7. J. C. Schouten, F. Takens, and C. M. van den Bleek, *Phys. Rev. E* **49** (1994) 126.
8. H. D. I. Abarbanel, *Analysis of Observed Chaotic Data*, (Springer Publ., New York, 1996).
9. L. M. Hively, P. C. Gailey, and V.A. Protopopescu, to appear in *Phys. Lett. A.*
10. L. M. Hively, P. C. Gailey, and V. A. Protopopescu, *to be published*.
11. V. E. Bondarenko, *Int. J. Bifur. Chaos* **7** (1997) 1133.
12. K. Lehnertz and C. E. Elger, *Phys. Rev. Lett.* **80** (1998) 5019.
13. L. D. Iasemidis and J. C. Sackellares, *Neuroscientist* **2** (1996) 118.
14. L. M. Hively, N. E. Clapp, C. S. Daw, W. F. Lawkins, and M. L. Eisenstadt, *ORNL/TM-12961* (Oak Ridge National Laboratory, Oak Ridge, TN) 1995.