ENHANCEMENTS IN EPILEPSY FOREWARNING VIA PHASE-SPACE DISSIMILARITY

(accepted for publication in *Journal of Clinical Neurophysiology*, August 2005)

L.M. Hively, V.A. Protopopescu, and N.B. Munro

Oak Ridge National Laboratory, Oak Ridge, TN 37831-6418

Corresponding author: L.M. Hively
Oak Ridge National Laboratory
P.O. Box 2008
Oak Ridge, TN 37831-6418
Phone: 865-574-7188
Fax: 865-576-5943

Running title: Enhancements in epilepsy forewarning

Enhancements in Epilepsy Forewarning Via Phase-Space Dissimilarity

We extend the recent application of phase space dissimilarity measures for scalp EEG data in two directions. First, we use a forewarning window of up to eight hours, thereby providing more forewarning time of the seizure event. This window was limited to a maximum of one hour in our previous work. Second, we combine information from two channels via a multi-channel phase-space to improve the quality and confidence limits of the forewarning. Combining these two enhancements, we obtain two-channel results that are superior to the single-channel ones.

***Index Terms*— dynamical systems, epileptic seizure forewarning, nonlinear analysis, phase-space dissimilarity measures, multi-channel analysis.**

Epilepsy afflicts about 1% of the world's population, or more than 50 million people. Typical seizures are not serious medical events in themselves, but may become life-threatening while the patient is pursuing hazardous activities (e.g., driving or swimming alone). However, extreme epileptic events require immediate medical intervention to avoid sudden unexplained death, which is characterized by fatal cardiac arrhythmia and/or breathing cessation, along with concomitant injuries. Two-thirds of patients have events that are controllable by anti-epileptic drugs, but the medications frequently have debilitating side effects (e.g., sleepiness, fuzzy thinking, and disorientation). Another 7-8% can be cured by epilepsy surgery, which may result in cognitive impairments. For the remaining 25% though, no available therapy is effective. Reliable seizure forewarning would allow preventive action, reduce morbidity and mortality, and improve patients' quality of life.

Early work on prediction of epileptic seizures began in the 1970s (Viglione and Walsh, 1975), expanding rapidly over the last decade, due to digital electroencephalographic (EEG) technology and advances in nonlinear dynamics (Eckmann and Ruelle, 1985; Jackson, 1989; Strogatz, 1994; Abarbanel, 1996; Kantz and Schreiber, 1999). Babloyantz and Destexhe(1986) and Babloyantz (1990a,b) suggested

that EEG data have noisy deterministic features that produce diverse behaviors, including chaos, although some investigators have challenged this idea (Ivanov et al., 1996; Jeong et al., 1999; Gribkov and Gribkova, 2000). The *Journal of Clinical Neurophysiology* published a recent focus issue (May 2001) on epilepsy prediction (Jerger et al., 2001; Lehnertz et al., 2001; Le Van Quyen et al., 2001; Osorio et al., 2001; Protopopescu et al., 2001; Savit et al., 2001; Sunderam et al., 2001). Litt and Echauz reviewed this research in May 2002, including time- and frequency-domain analysis, nonlinear dynamics and chaos, as well as neural networks and other artificial intelligence approaches. *IEEE Transactions on Biomedical Engineering* published a focus issue (May 2003) on prediction of epilepsy (Chavez et al., 2003; D'Alessandro et al., 2003; Hively and Protopopescu, 2003; Iasemidis, 2003; Iasemidis et al., 2003; Lopes da Silva et al., 2003; McSharry et al., 2003; Notley and Elliott, 2003; Paul et al., 2003; Rieke et al., 2003; Slutzky et al., 2003; Witte et al., 2003). Typical measures for prediction include the largest Lyapunov exponent (Iasemidis, 2003), synchrony (Chavez et al., 2003), correlation integral (Notley and Elliott, 2003), and various time- and frequency-domain features of EEG energy (D'Alessandro et al., 2003). These results are mostly based on analysis of *intracranial* EEG.

We use *scalp* EEG, which is non-invasive, but has resisted previous analyses due to signal attenuation through bone and soft tissue, and contamination from eye blinks and other muscular artifacts. Our initial analysis of scalp EEG (Hively et al., 1995) used traditional nonlinear measures (TNM), such as Kolmogorov entropy and correlation dimension, yielding inconsistent forewarning. In our subsequent work (Hively et al., 1999; Hively, Clapp et al., 2000; Hively, Gailey et al., 2000; Hively, Protopopescu et al., 2000; Gailey et al., 2001; Protopopescu et al., 2001; Hively and Protopopescu, 2003), we developed and implemented the phase-space dissimilarity measures (PSDM). We also showed by direct comparison that these measures provide consistently better forewarning than TNM, independent of patients' age, sex, event onset time, pre-event activity, or awake- versus asleep-state baseline state. The improved discrimination by PSDM is a direct result of their definition and calculation by summing the absolute value of differences, while TNM are based on a difference of averages.

Previously, we limited the forewarning analysis to no more than one hour before the clinically characterized seizure. This forewarning window was motivated by two reasons, namely: (i) we wanted to see whether any forewarning can be obtained at all; and (ii) forewarnings longer than one hour were generally considered too long and/or too uncertain to be of clinical value. However, recent work by Litt and Echauz (2002) demonstrated forewarning of several hours before the epileptic events. Consequently, the present work extends the time window to the full length of the data, yielding forewarning of ≤6 hours, which is consistent with Litt and Echauz (2002) .

In our previous work (Hively et al., 1999; Hively, Clapp et al., 2000; Hively, Gailey et al., 2000; Hively, Protopopescu et al., 2000; Gailey et al., 2001; Protopopescu et al., 2001; Hively and Protopopescu, 2003), we used forewarning indication (or lack thereof in non-event datasets) in any one channel, regardless of the results in other channels. Specifically, instances of a true positive (true negative) in one channel are sometimes accompanied by a false negative (or false positive) in other channel(s) for an event (non-event) dataset. Consequently, the any-one channel criterion is rather weak and clearly needs improvement. In the present work, we abandon the any-one channel criterion, and instead use simultaneous information from a combination of two fixed channels with a resultant forewarning that is consistently more robust than the forewarning from either of the individual channels alone. This result is consistent with our intuitive expectation that the use of two channels should increase the amount of information in comparison to the single-channel analysis, thereby providing a stronger forewarning for all patients.

The remainder of this paper is organized as follows. The Methods section describes the data acquisition, artifact filtering, PSDM methodology, and approach for channel-consistent forewarning. The Results section explains our analysis and presents the results. The Discussion section presents our concluding remarks.

METHODS

We retrospectively analyzed 60 scalp EEG datasets: 40 with at least one electrographic temporal-

lobe event, plus 20 datasets without epileptic events as controls. The data were obtained from 41 different patients with ages between 4 and 57 years. Half of the patients have only a single dataset. The other 30 datasets are from 11 different patients with multiple datasets: 7 patients with 2 datasets, one patient with 3 datasets, 2 patients with 4 datasets, and one patient with 5 datasets. Our analysis begins at the start of each dataset, and ends with the last full window of data or with the window that includes the first event (whichever comes first). The dataset lengths are between 5,016 s (1 h and 23 min) and 29,656 s (8 h and 14 min), with an average of 15,668 s (4 h and 21 min). The cumulative period of the analysis is 940,104 s (261 h and 8 min). If we split these 60 datasets into equally-sized training and test sets, then inadequate statistics would result. Specifically, the resultant training (and test) sets would then have only ten non-event datasets. Moreover, only fifteen datasets from the patients with multiple datasets would be available for the channel-consistency analysis for the training (and test) sets. Consequently, we use all of the datasets as a training set, which clearly limits the strength of our conclusions.

Acquisition and preparation of EEG data

EEG data were acquired from each patient over several days to two weeks under standard clinical protocols. Recordings came from 32-channel instruments (Nicolet-BMSI, Madison, Wisconsin) with 19 scalp electrodes in the International 10−20 system of placement (see Fig. 1). Each channel was amplified separately, band-pass filtered between 0.5−99 Hz, digitized at 250Hz, and stored on VHS tape. The physicians independently selected data from this archive, as “representative” of the patients’ condition, without regard to the presence (or absence) of artifacts or other noise. To ensure adequate data for our analysis, we further require that: (i) event datasets include a baseline period of at least two hours plus the entire event segment, and (ii) non-event datasets be longer than one hour. The physician’s characterization covered only the dataset, which represents a fraction of the total monitoring periods, and in principle limits the generality of our results. Analysis of the full monitoring periods is beyond the scope of this work. The characterization of each dataset includes: patient’s sex and age; start and stop time of the data recording; seizure event onset time; seizure type; and the patient’s activity at the start of

the recording and immediately prior to the event, as shown in Table 1 of our earlier work (Protopopescu et al., 2001).

We first divide the stream of time-serial data into contiguous, non-overlapping windows (cutsets) of N data points. The artifact signal is removed from each cutset with a novel zero-phase quadratic filter (Hively et al., 1995), using a moving window of data points, e_i, with the same number of data points, w, on either side of a central point. We fit a quadratic curve in the least-squares sense over this window, taking the central point of the fit as the low-frequency artifact, f_i. The residual value, $g_i = e_i - f_i$, has essentially no low-frequency artifact activity.

Phase-space methodology

We convert each artifact-filtered point into a discrete symbol, s_i, as one of S different integers, $0 \leq s_i \leq S - 1$. We use contiguous, non-overlapping partitions to obtain uniform symbols between the minimum, g_{min}, and maximum, g_{max}, in the data of the first base case cutset with $s_i = INT[S(g_i - g_{min})/(g_{max} - g_{min})]$. The function INT converts a decimal number to the next lower integer, e.g., $INT(3.14) = 3$. Alternatively, one can use equiprobable symbols, by ordering all N baseline data points from the smallest to largest value. The resultant symbolization is: $s_i = 0$ for $g_{min} \leq g_i < g_1$; $s_i = 1$ for $g_1 \leq g_i < g_2$; ...; $s_i = S - 2$ for $g_{S-2} \leq g_i < g_{S-1}$; and $s_i = S - 1$ for $g_{S-1} \leq g_i \leq g_{max}$. The first N/S of these ordered data values correspond to the first symbol, 0, by an appropriate choice of g_1. Ordered data values $(N/S)+1$ through $2N/S$ correspond to the second symbol, 1, and so on up to the last symbol, $S - 1$. Equiprobable symbols have non-uniform partitions in signal amplitude with dynamical structure arising only from the phase space (PS) reconstruction of Eqs. (1)-(2). Large negative and/or positive values of g_i have little effect on equiprobable symbolization, but significantly change the partitions for uniform symbols.

We assume that the complex, high-dimensional brain dynamics are confined to a bounded, low-dimensional region, called an “attractor.” Then, the symbolized data can be converted into a phase space representation by standard reconstruction of the dynamics via time-delay vectors (Eckmann and Ruelle, 1985). The single-channel form is:

(1) $y(i) = [s_i, s_{i+\lambda}, \ldots, s_{i+(d-1)\lambda}]$

The symbolization discretizes the PS representation into S^d bins to extract event forewarning on the basis of a local time delay, λ, and dimensionality, d. In addition, information exchange in the brain connects local processes, implying that a multi-channel PS vector of C channels (Boccaletti et al., 2002) could contain more information. Here, the underlying assumption is that additional channels add more information about the dynamics that are connected via the overall process. The resultant multi-channel PS form is:

(2) $y(i) = [s(1)_i, s(1)_{i+\lambda}, \ldots, s(1)_{i+(d-1)\lambda}, \ldots, s(C)_i, s(C)_{i+\lambda}, \ldots, s(C)_{i+(d-1)\lambda}]$,

where, $s(k)$ denotes symbols from the kth channel, $1 \le k \le C$. Symbolization divides the multi-channel PS into S^{Cd} bins. We count the number of points that occur in each PS bin, thus forming a discrete DF that accounts for the geometry and visitation frequency on the attractor. We denote the population of the ith bin of the DF, Q_i, for the base case, and R_i for a test case, respectively. An (un)changing DF indicates (un)altered dynamics.

We compare the test case to the base case via dissimilarity measures (DM) between Q_i and R_i (Protopopescu et al., 2001):

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

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

These sums run over all populated PS cells. The χ^2 measure is a powerful, robust, and widely used comparison between observed and expected frequencies. The χ^2 is a *relative* measure (Hively et al., 1999) of dissimilarity in this context, rather than a mathematical distance or an unbiased statistic for accepting or rejecting a null statistical hypothesis. The L_1 distance is the natural metric for DFs by its direct relation to the total invariant measure on the attractor and defines a *bona fide* distance. Consistent calculations require the same number of points in both Q_i and R_i, identically sampled; otherwise the DFs must be properly rescaled. We also capture the dynamical flow (Abarbanel, 1996) between PS states,

$y(i) \rightarrow y(i+1)$, as a connected PS (CPS) vector, $Y(i) = [y(i), y(i+1)]$, with corresponding DFs and dissimilarity measures (denoted below by the subscript, c) that have the same form as the PS measures (Protopopescu et al., 2001; Hively and Protopopescu, 2003).

We use the DFs from the first B cutsets to represent non-seizure (baseline) dynamics, with B=10 to capture sufficient variability. The baseline DFs are exhaustively compared to one another in pair-wise fashion via Eqs. (3) - (4) to obtain the mean baseline dissimilarity, $\underline{V}$, as well as a corresponding standard deviation, σ, for each DM from the set, $V = \{L, L_c, \chi^2, \chi_c^2\}$. The disparate range and variability of these measures are difficult to interpret, particularly for noisy data, so we need a consistent criterion for comparison. For this purpose, we renormalize the dissimilarity measures (Protopopescu et al., 2001; Hively and Protopopescu, 2003) by comparing each of the B baseline cutsets to each (ith) test case cutset, and then computing the corresponding average dissimilarity value, V_i, of the ith cutset. The renormalized dissimilarity is: $U(V) = |V_i - \underline{V}|/\sigma$, as the number of standard deviations that the test case deviates from the baseline mean. We use these renormalized dissimilarity measures to test for statistically significant change.

Data analysis

The analysis starts at the beginning of the dataset and proceeds forward in time until a forewarning occurs, as defined next. We define a true positive (TP = 1 in the sum below) as a forewarning in any one channel when a number of simultaneous PSDM, N_{SIM}, exceed a threshold, $U \geq U_C$, for a number of sequential occurrences, N_{OCC}, within a preset forewarning window, 1 minute $\leq T_1 \leq T_{SZ} - T_{FW} \leq T_2$, before the seizure onset time, T_{SZ}. The corresponding forewarning time is T_{FW}. The value of T_1 was chosen, based on input from a physician collaborator that even with one minute of forewarning, useful things could be done to help the patient medically[1]. A false positive (FP) occurs if $T_{SZ} - T_{FW} < T_1$ or $T_{SZ} - T_{FW} > T_2$ in an event dataset, or for a forewarning in a non-event dataset. A true

[1] Private Communication from M. L. Eisenstadt, M.D., Ph.D.

negative (TN = 1) corresponds to no forewarning in a non-event dataset. No forewarning in an event dataset is a false negative (FN). The number of true instances of channel-consistent forewarning (Hively and Protopopescu, 2003) for the jth channel of the kth patient is $T_{jk} = \Sigma_i [TP_{ijk} + TN_{ijk}]$, where the sum over datasets runs from i=1 to $M(k)$ = number of datasets for the kth patient. The occurrence of $T_{jk} \geq 2$ indicates consistency in more than one dataset for the same patient, while $T_{jk} \leq 1$ means that the jth channel provides no such consistency. The best channel consistency is $c_k = \max(T_{jk})$, for $T_{jk} \geq 2$ and k fixed; $c_k = 0$, if $T_{jk} \leq 1$; $c_k = 1$ for patients with only one dataset. The channel-consistent total-true rate then becomes $f_T = [\Sigma_k c_k]/[\Sigma_k M(k)]$. Here, k runs over all P of the patients, which weights each dataset equally.

To maximize f_T, we searched over parameters in the following ranges: $5\,000 \leq N \leq 100\,000$, $2 \leq w \leq 1\,000$, $2 \leq S \leq 200$, $2 \leq d \leq 26$ (single channel via Eq. 1) or $2 \leq d \leq 13$ (two-channel via Eq. 2), and $1 \leq \lambda \leq 100$. We use double-precision base-S arithmetic to represent the PS state, thus limiting the combination of values of S and d. A huge reduction in the computational effort is obtained by varying one parameter at a time, while holding the others fixed. This approach avoids an exhaustive multi-dimensional search at the expense of sub-optimal, but quite acceptable results (Protopopescu et al., 2001). The forewarning analysis of the previous paragraph exhaustively searches over other parameters: $1 \leq N_{SIM} \leq 4$, $1 \leq N_{OCC} \leq 100$, $0.001 \leq U_C \leq 2$. After determining the best parameters, the analysis is much faster than the data-recording time.

RESULTS

As explained in the introduction, our previous work (Protopopescu et al, 2001; Hively and Protopopescu, 2003) restricted the forewarning window, $T_2 \leq 1$ hour, for artifact-filtered, monopolar data. Here, we first examine the effect of opening the forewarning window to $T_2 \leq 8$ hours, for combinations of unfiltered and artifact-filtered, monopolar and bipolar data. Figure 1 shows the electrode locations. Table 1 (rows 1-8) shows monopolar results (M in column 1) with artifact filtering on (column 2) for various values of T_2 (column 3), and corresponding values of N_{SIM} (column 4), N_{OCC} (column 5),

and U_C (column 6). This analysis uses the best set of parameters (w=54, N=22000, d=2, S=20, λ=17) from our most recent work (Hively and Protopopescu, 2003) for equiprobable symbols and single-channel analysis via Eq. (1). Statistical measures of the forewarning times (in seconds) include the minimum (MIN in column 7), maximum (MAX in column 8), average (AVG in column 9), and sample standard deviation (SD in column 10). Columns 11-17 display the distribution of T_{FW} in one hour increments, from ≤1 hour, 1-2 hours, and so on up to 6-7 hours. The magnitude of f_T depends in this case on the (weak) any-one-channel criterion, and always rises monotonically as T_2 increases. Consequently, we cite only changes in f_T for each increase in T_2, $\Delta f_T = x/60$. For monopolar analysis with artifact filtering, we obtain x = 3 and 2 as the maximum forewarning time increases from $T_2 \leq 1$ hour to $T_2 \leq 2$ hours, and from $T_2 \leq 2$ hour to $T_2 > 2$ hours, respectively. Rows 9-16 show monopolar results without artifact filtering. Here, we obtain x = 12, 1, 1, 1 when the maximum forewarning time increases from $T_2 \leq 1$ hour to $T_2 \leq 2$ hours, $T_2 \leq 2$ hours to $T_2 \leq 3$ hours, $T_2 \leq 3$ hours to $T_2 \leq 4$ hours, and from $T_2 \leq 4$ hours to $T_2 > 4$ hours, respectively. These f_T-values are smaller than the artifact-filtered counter-parts for $T_2 \leq 4$ hours, and then are the same for $T_2 > 4$ hours. However, the distribution of T_{FW} without artifact filtering extends beyond 6 hours, which is not sufficiently conclusive for clinical use. Rows 17-24 show results with artifact filtering for 18 bipolar channels (B in column 1). A channel in this case is the difference between adjacent monopolar channels, removing most locally-measured artifacts, while retaining a differential measure of the brain dynamics. Table 2 includes the definitions of the bipolar channels. Here, we obtain x = 10 and 1 as the maximum forewarning time increases from $T_2 \leq 1$ hour to $T_2 \leq 2$ hours, and from $T_2 \leq 2$ hour to $T_2 > 2$ hours, respectively. These f_T values are below the counterpart monopolar results only for $T_2 \leq 1$ hour, and then are consistently larger for $T_2 > 1$ hour. Rows 25 through 32 show forewarning for the bipolar montage without artifact filtering. Here, we obtain values of x = 3 and 2 as the maximum forewarning time increases from $T_2 \leq 1$ hour to $T_2 \leq 2$ hours, and from $T_2 \leq 2$ hour to $T_2 > 2$ hours, respectively. These f_T values fall below the artifact-filtered bipolar results only for $1 \leq T_2 < 2$ hours. We note that $N_{SIM} = 4$ occurs for bipolar channels with and without artifact filtering for $T_2 > 4$ hours, which is a very robust

result that is not consistently displayed by the monopolar results. These results show the superiority of artifact-filtered bipolar EEG, with a smallest maximum forewarning of 5.32 hours (19,184 s), consistent with recent work by Litt and Echauz, 2002).

Another extension of the methodology is the multi-channel form of Eq. (2). Based on the results of the previous paragraph (Table 1), we analyze two artifact-filtered, bipolar channels. An exhaustive analysis of all possible pairs of eighteen bipolar channels would involve 153 ($18 \times 17/2$) unique combinations, which is beyond the scope of the present work. Instead, we restrict the analysis to symmetric pairs of bipolar channels, which are expected to contain complementary information. Table 2 shows the results for the single-channel analyses of Table 1, in comparison with dual-channel analysis for equivalent parameters (w=54, N=22000, d=1, S=20) with a forewarning window of one hour. Channel consistency is not required for this portion of the analysis, since the channel is fixed in each case. No value of λ is needed for the dual-channel analysis, because Eq. (2) for d=1 reduces to $y(i) = [s(1)_i, s(2)_i]$. Larger values of T_2 increase the total true rate, as shown in Table 1. However, the point of this additional analysis is the relative improvement in f_T from single-channel to dual-channel analysis. As expected, the dual channel analysis yields a larger total true rate than either of the single-channel analyses alone. The extreme frontal, bipolar channels have the best total true rates, even though these EEG data are for temporal-lobe epileptic events. The reason for this puzzling result may be that channels closer to the temporal lobes are too sensitive, while those somewhat farther away give appropriate sensitivity for forewarning. We note that the frontal (forehead) channels have minimal hair, potentially enabling electrode placement by the patient or a care-giver for long-term, non-clinical, ambulatory monitoring.

A further improvement is optimization of the parameters for the artifact-filtered, bipolar pair with the best total true in Table 2 (C1-C9). A systematic search (Hively and Protopopescu, 2003) yields a result that has one false-negative and three false-positive indications. This result corresponds to an average rate of 0.011 false positives per hour, or equivalently a mean time between false positives of 87 hours. This FP rate is more than an order of magnitude below the maximum acceptable rate of 0.15FP/hour, as recommended by Winterhalder et al. (2003). We obtain this result for uniform symbols

with w=280, N=19316, d=1, and S=55. (Equiprobable symbols give even lower values of f_T in this case.) This instance has no λ, because Eq. (2) for d=1 reduces to $y(i) = [s(C1)_i, s(C9)_i]$. While this result is an improvement over our earlier work (Hively and Protopopescu, 2003), the choice of parameters is not robust, because a unit change in any of these parameters lowers the total true rate. Thus, improvement in the forewarning robustness is an important focus for future work. Figure 2 shows the corresponding distribution of forewarning times. We note that the cumulative distribution function of forewarning times in Fig. 2 rises roughly linearly from 0.5h to 3.5h, while encompassing 90% of the forewarnings. Figure 3 shows an example of the χ^2 dissimilarity measures for the analysis of channel C1 only, channel C9 only, and dual-channel analysis for C1 and C9. This figure shows that the dual-channel-result extracts the better forewarning from channel C9 rather than that from C1, as expected.

DISCUSSION

We deem these results as encouraging and worthy of continuing development despite several limitations, which we discuss next. First, we use all of the data as a training set, limiting the strength of our conclusions. However, the alternative would involve equally-sized training and test sets. The resultant training (and test) sets would then have only ten non-event datasets and fifteen datasets for channel-consistency, which would result in inadequate statistics. Second, we analyze EEG data from a controlled clinical setting, rather than an uncontrolled, ambulatory environment. Third, our results depend on careful adjustment of the analysis parameters for the best total-true rate. Fourth, we analyze only physician-selected portions of the patient data, rather than the full monitoring period. Fifth, these results are for only sixty datasets (forty with epileptic events) for a single seizure type (temporal lobe). Much more data (hundreds of datasets spanning >1000 h) are needed for the proper choice of the analysis parameters as part of a robust and conclusive statistical validation. These data requirements are far beyond our present capabilities, and almost everyone else's, based on present publications. Sixth, we have not performed prospective analysis of long-term continuous data, which is the acid test for any predictive approach. Finally, the present analyst-intensive methodology uses retrospective analysis of archival EEG on a desktop computer. Ambulatory forewarning requires analyst-independent, prospective analysis of real-

time EEG on a portable device. Clearly, much work remains to address these issues. Thus, we view the importance of this work as an example of the overall potential of the methodology, rather than the specific results.

Acknowledgments: This work was supported in part by the U.S. Department of Energy's Division of Materials Science and Engineering under the Office of Basic Energy Science. We thank Nicolet Biomedical Inc., Madison, Wisconsin (now part of ViaSys Healthcare Inc.) for providing and characterizing the EEG data under a Cooperative Research and Development Agreement with Oak Ridge National Laboratory, which is managed for the U.S. Department of Energy by UT-Battelle, LLC, under Contract No. DE-AC05-00OR22725.

REFERENCES

Abarbanel HDI. *Analysis of observed chaotic data*. New York: Springer Verlag, 1996.

Babloyantz A. Chaotic dynamics in brain activity. In: Basar E, ed. *Chaos in brain function*. New York: Springer Verlag, 1990a;42-8.

Babloyantz A. Estimation of correlation dimension. In: Basar E, ed. *Chaos in brain function*. New York: Springer Verlag, 1990b;83-91.

Babloyantz A, Destexhe A. Low dimensional chaos in an instance of epilepsy. *Proc Natl Acad Sci USA* 1986;83:3513-7.

Boccaletti S, Valladares DL, Pecora LM, Geffert HP, Carroll T. Reconstructing embedded spaces of coupled dynamical systems from multivariate data. *Phys Rev E* 2002;65:035204.

Chavez M, Le Van Quyen M, Navarro V, Baulac M, Martinerie J. Spatio-temporal dynamics prior to neocortical seizures: amplitude versus phase couplings. *IEEE Trans Biomed Eng* 2003;50**:**537-9.

D'Alessandro M, Esteller R, Vachtsevanos G, Hinson A, Echauz J, Litt B. Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. *IEEE Trans Biomed Eng* 2003;50**:**603-15.

Eckmann JP, Ruelle D. Ergodic theory of chaos and strange attractors. *Rev Mod Phys* 1985;57: 617-655.

Gailey PC, Hively LM, Protopopescu VA. Robust detection of dynamical change in EEG. In: Ditto WL, et al., eds. *Proceedings of the 5th Experimental Chaos Conference, Orlando, Florida, 28 June-1 July: 1999.* Singapore: World Scientific, 2001.

Gribkov V, Gribkova D. Learning dynamics from nonstationary time series: analysis of electroencephalograms. *Phys Rev E* 2000;61:6538-45.

Hively LM, Clapp NE, Daw CS, Lawkins WF. *Nonlinear analysis of EEG for epileptic seizures.* ORNL/TM-12961. Oak Ridge, TN: Oak Ridge National Laboratory, 1995.

Hively LM, Clapp NE, Protopopescu VA, Joseph J, Merican CE, Lucht T. *Epileptic seizure forewarning by nonlinear techniques.* ORNL/TM-2000/333. Oak Ridge, TN: Oak Ridge National Laboratory, 2000.

Hively LM, Gailey PC, Protopopescu VA. Detecting dynamical change in nonlinear time series. *Phys Lett A*, 1999;258**:**103–14.

Hively LM, Gailey PC, Protopopescu VA. Sensitive measures of condition change in EEG data. In: Lehnertz K, et al., eds. *Proceedings of the International Workshop "Chaos in Brain?," Bonn, Germany, 10-12 March:1999.* Singapore: World Scientific, 2000;333-6.

Hively LM, Protopopescu VA. Channel-consistent forewarning of epileptic events from scalp EEG. *IEEE Trans Biomed Eng* 2003;50:584-93.

Hively LM, Protopopescu VA, Gailey PC. Timely detection of dynamical change in scalp EEG signals. *Chaos* 2000;10:864-75.

Iasemidis LD. Epileptic seizure prediction and control (Invited Paper). *IEEE Trans Biomed Eng* 2003;50**:** 549-58.

Iasemidis LD, Shiau D-S, Chaovalitwongse W, Sackellares JC, Pardalos PM, Principe JC, Carney PR, Prasad A, Veeramani B, Tsakalis K. Adaptive epileptic seizure prediction system. *IEEE Trans Biomed Eng* 2003;50:616-27.

Ivanov DK, Posch HA, Stumpf C. Statistical measures derived from the correlation integrals in physiological time series. *Chaos* 1996;6:243-53.

Jackson EA. *Perspectives of nonlinear dynamics*. Vol. 1. Cambridge: Cambridge University Press, 1989.

Jackson EA. *Perspectives of nonlinear dynamics*. Vol. 2. Cambridge: Cambridge University Press, 1990.

Jeong J, Kim MS, Kim SY. Test for low-dimensional determinism in EEG. *Phys Rev E* 1999;60**:**831-7.

Jerger KK, Netoff TI, Francis JT, Sauer T, Pecora L, Weinstein SL, Schiff SJ. Early seizure detection. *J Clin Neurophysiol* 2001;18:259-68.

Kantz H, Schreiber T. *Nonlinear time series analysis*. Cambridge: Cambridge University Press, 1999.

Lehnertz K, Andrzejak RG, Arnhold J, Kreuz T, Mormann F, Reike C, Widman G, Elger CE. Nonlinear EEG analysis in epilepsy: its possible use for interictal focus localization, seizure anticipation, and prevention. *J Clin Neurophysiol* 2001;18:209-22.

Le Van Quyen M, Martinerie J, Navarro V, Baulac M, Varela FJ. Characterizing neurodynamic changes before seizures. *J Clin Neurophysiol* 2001;18:191-208.

Litt B, Echauz J. Prediction of epileptic seizures. *Lancet Neurology* 2002;1:22-30.

Lopes da Silva FH, Blanes W, Kalitizin SN, Parra J, Suffczynski P, Velis DN. Dynamical diseases of brain systems: different routes to epileptic seizures (Invited Paper). *IEEE Trans Biomed Eng* 2003;50:540-8.

McSharry PE, Smith LA, Tarassenko L. Comparison of predictability of epileptic seizures by a linear and a nonlinear method. *IEEE Trans Biomed Eng* 2003;50:628-33.

Notley SV, Elliott SJ. Efficient estimation of a time-varying dimension parameter and its application to EEG analysis. *IEEE Trans Biomed Eng* 2003;50:594-602.

Osorio I, Harrison MAF, Lai Y-C, Frei MG. Observations on the application of the correlation dimension and correlation integral to the prediction of seizures. *J Clin Neurophysiol* 2001;18:269-74.

Paul JS, Patel CB, Al-Nahash H, Zhang N, Ziai WC, Mirski MA, Sherman DL. Prediction of PTZ-induced seizures using wavelet-based residual entropy of cortical and subcortical field potentials. *IEEE Trans Biomed Eng* 2003;50**:**640-8.

Protopopescu VA, Hively LM, Gailey PC. Epileptic seizure forewarning from scalp EEG. *J Clin Neurophysiol* 2001;18:223-45.

Rieke C, Mormann F, Andrzejak RG, Kreuz T, David P, Elger CE, Lehnertz K. Discerning nonstationarity from nonlinearity in seizure-free and preseizure EEG recordings from epilepsy patients. *IEEE Trans Biomed Eng* 2003;50**:**634-9.

Savit R, Li D, Zhou W, Drury I. Understanding dynamical state changes in temporal lobe epilepsy. *J Clin Neurophysiol* 2001;18:246-58.

Slutzky MW, Cvitanovic P, Mogul DJ. Manipulating epileptiform bursting in the rat hippocampus using chaos control and adaptive techniques. *IEEE Trans Biomed Eng* 2003;50**:**559-70.

Strogatz SH. *Nonlinear dynamics and chaos.* Reading: Addison Wesley, 1994.

Sunderam S, Osorio I, Frei MG, Watkins JF III. Stochastic modeling and prediction of experimental seizures in Sprague-Dawley rats. *J Clin Neurophysiol* 2001;18:275-82.

Viglione SS, Walsh GO. Proceedings: epileptic seizure prediction. *Electroencephalogr Clin Neurophysiol* 1975;39:435-6.

Winterhalder M, Maiwald T, Voss HU, Aschenbrenner-Scheibe R, Timmer J, Schulze-Bonhage A. The seizure prediction characteristic: a general framework to access and compare seizure prediction methods. *Epilepsy Behav* 2003;4:318-25.

Witte H, Iasemidis LD, Litt B. Special issue on epileptic seizure prediction. *IEEE Trans Biomed Eng* 2003;50:537-9.

Table 1: Largest f_T for monopolar (M) and bipolar (B) montage with artifact filtering (AF) on or off

	MB	AF	T_2(h)	N_{SIM}	N_{OCC}	U_C	T_{FW}(s) MIN	MAX	AVG	SD	Distribution of T_{FW} 1h	2h	3h	4h	5h	6h	7h
1)	M	on	1	4	54	0.012	60	3588	2095	905	77						
2)	M	on	2	4	41	0.017	68	7164	4229	1945	63	108					
3)	M	on	3	≥3	43	0.073	76	10772	6153	2739	60	126	118				
4)	M	on	4	4	41	0.016	68	14396	7196	3370	72	149	159	66			
5)	M	on	5	4	41	0.016	68	17652	7614	3772	72	149	159	66	23		
6)	M	on	6	4	41	0.016	68	20512	7753	3954	74	151	161	68	23	6	
7)	M	on	7	4	41	0.016	68	20512	7753	3954	74	151	161	68	23	6	
8)	M	on	8	4	41	0.016	68	20512	7753	3954	74	151	161	68	23	6	
9)	M	off	1	≥3	53	0.083	140	3576	2096	1020	53						
10)	M	off	2	4	54	0.011	76	7196	4115	2025	64	103					
11)	M	off	3	≥3	47	0.066	136	10720	6271	2817	56	132	125				
12)	M	off	4	4	46	0.014	120	14268	7424	3639	57	139	114	75			
13)	M	off	5	3	2	0.015	316	17776	8619	4658	28	42	41	35	22		
14)	M	off	6	3	2	0.015	316	20944	9281	5212	28	42	41	35	22	11	
15)	M	off	7	3	2	0.015	316	21876	9351	5282	28	42	41	35	22	11	1
16)	M	off	8	3	2	0.015	316	21876	9351	5282	28	42	41	35	22	11	1
17)	B	on	1	≥3	50	0.063	72	3552	1883	1016	72						
18)	B	on	2	≥3	49	0.062	132	7184	4285	2134	66	107					
19)	B	on	3	≥3	47	0.062	176	10744	6016	2725	66	118	119				
20)	B	on	4	4	34	0.030	76	14380	7043	3425	54	131	109	53			
21)	B	on	5	4	34	0.030	76	17828	7575	3924	57	132	110	54	24		
22)	B	on	6	4	34	0.030	76	19184	7666	4040	57	132	110	54	24	3	
23)	B	on	7	4	34	0.030	76	19184	7666	4040	57	132	110	54	24	3	
24)	B	on	8	4	34	0.030	76	19184	7666	4040	57	132	110	54	24	3	
25)	B	off	1	4	56	0.012	144	3552	2211	938	72						
26)	B	off	2	4	56	0.013	104	7196	4171	1825	71	114					
27)	B	off	3	4	56	0.013	104	10572	5460	2514	68	120	75				
28)	B	off	4	4	34	0.027	76	14368	7510	3405	43	136	117	66			
29)	B	off	5	4	34	0.027	76	17740	7974	3824	46	139	119	67	24		
30)	B	off	6	4	34	0.027	76	20680	8173	4077	46	139	119	67	24	7	
31)	B	off	7	4	34	0.027	76	20680	8173	4077	46	139	119	67	24	7	
32)	B	off	8	4	34	0.027	76	20680	8173	4077	46	139	119	67	24	7	

Table 2: Total true values (f_T) for artifact-filtered bipolar analysis with T_1 = 1 hour

Bipolar channels	First Channel	Second Channel	Dual Channel
C01 = F7 – FP1& C09 = F8 – FP2	f_T(C01) =18/60	f_T(C09) =12/60	f_T(C01,C09) = 27/60
C02 = T3 – F7 & C10 = T4 – C8	f_T(C02) =13/60	f_T(C10) =16/60	f_T(C02,C10) = 22/60
C03 = T5 – T3 & C11 = T6 – T4	f_T(C03) =17/60	f_T(C11) =18/60	f_T(C03,C11) = 21/60
C04 = O1 – T5 & C12 = O2 – T6	f_T(C04) =15/60	f_T(C12) =09/60	f_T(C04,C12) = 21/60
C05 = F3 – FP1& C13 = F4 – FP2	f_T(C05) =22/60	f_T(C13) =21/60	f_T(C05,C13) = 26/60
C06 = C3 – F3 & C14 = C4 – F4	f_T(C06) =16/60	f_T(C14) =17/60	f_T(C06,C14) = 25/60
C07 = P3 – C3 & C15 = P4 – C4	f_T(C07) =15/60	f_T(C15) =15/60	f_T(C07,C15) = 22/60
C08 = O1 – P3 & C16 = O2 – P4	f_T(C08) =15/60	f_T(C16) =10/60	f_T(C08,C16) = 22/60
C17 = CZ – FZ & C18 = PZ – CZ	f_T(C17) =16/60	f_T(C18) =13/60	f_T(C17,C18) = 23/60

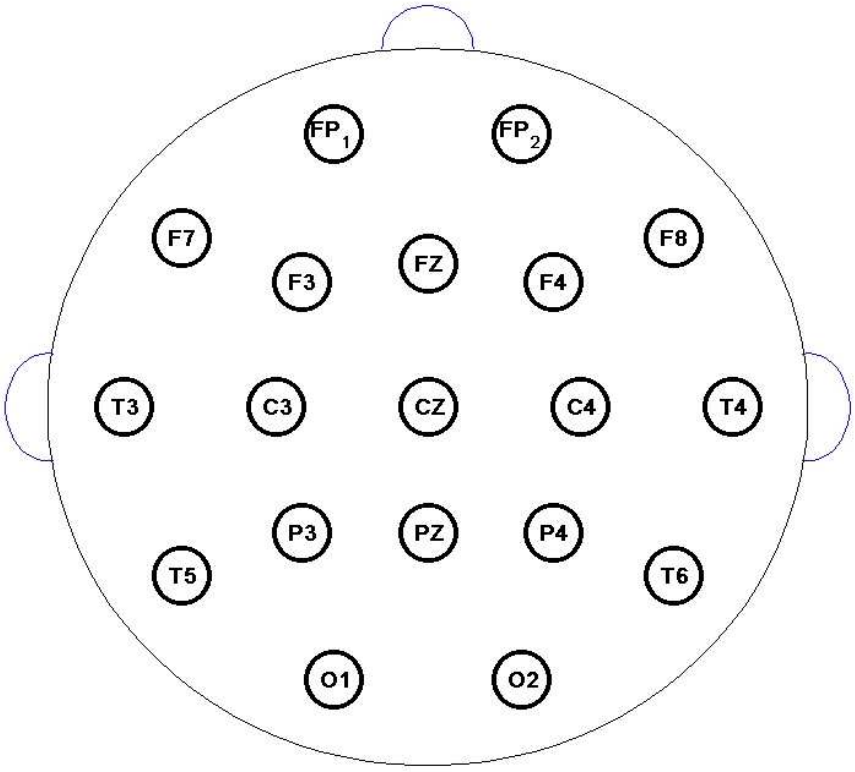

Fig. 1. Layout of the EEG scalp electrodes in the standard 10-20 montage, looking at the patient's head from above. The bipolar montage uses signal differences from adjacent electrodes, such as $C1 = F_7 - FP_1$ and $C_9 = F_8 - FP_2$, which are the symmetric channels, which are chosen for the multi-channel analysis via Eq. (2).

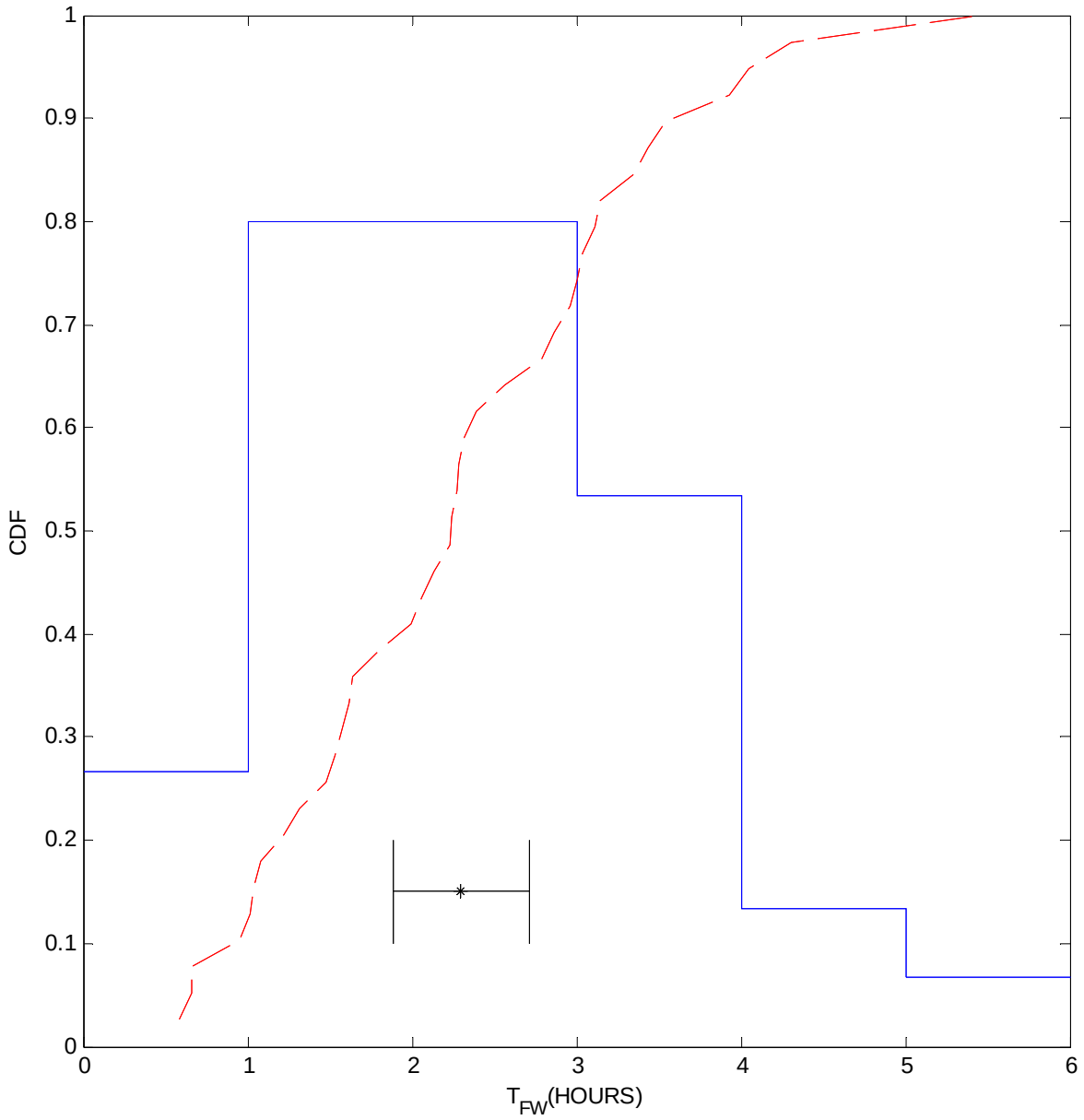

Figure 2: Distribution of forewarning time (T_{FW}) for the dual-channel analysis of bipolar scalp EEG with f_T = 56/60. The corresponding sensitivity, TP/(TP + FN), is 39/40; the specificity, TN/(TN + FP), is 17/20. The stair-step (solid) plot indicates the relative occurrence frequency (arbitrary units) of T_{FW} versus time intervals in one-hour increments, beginning at 0-1 hour, and ending with 5-6 hours. The monotonically increasing (dashed) curve is the cumulative distribution function (CDF) of T_{FW} versus time. The H-bar with the star in the middle indicates the mean value of T_{FW} (2.3 hours) and the standard deviation of the mean.

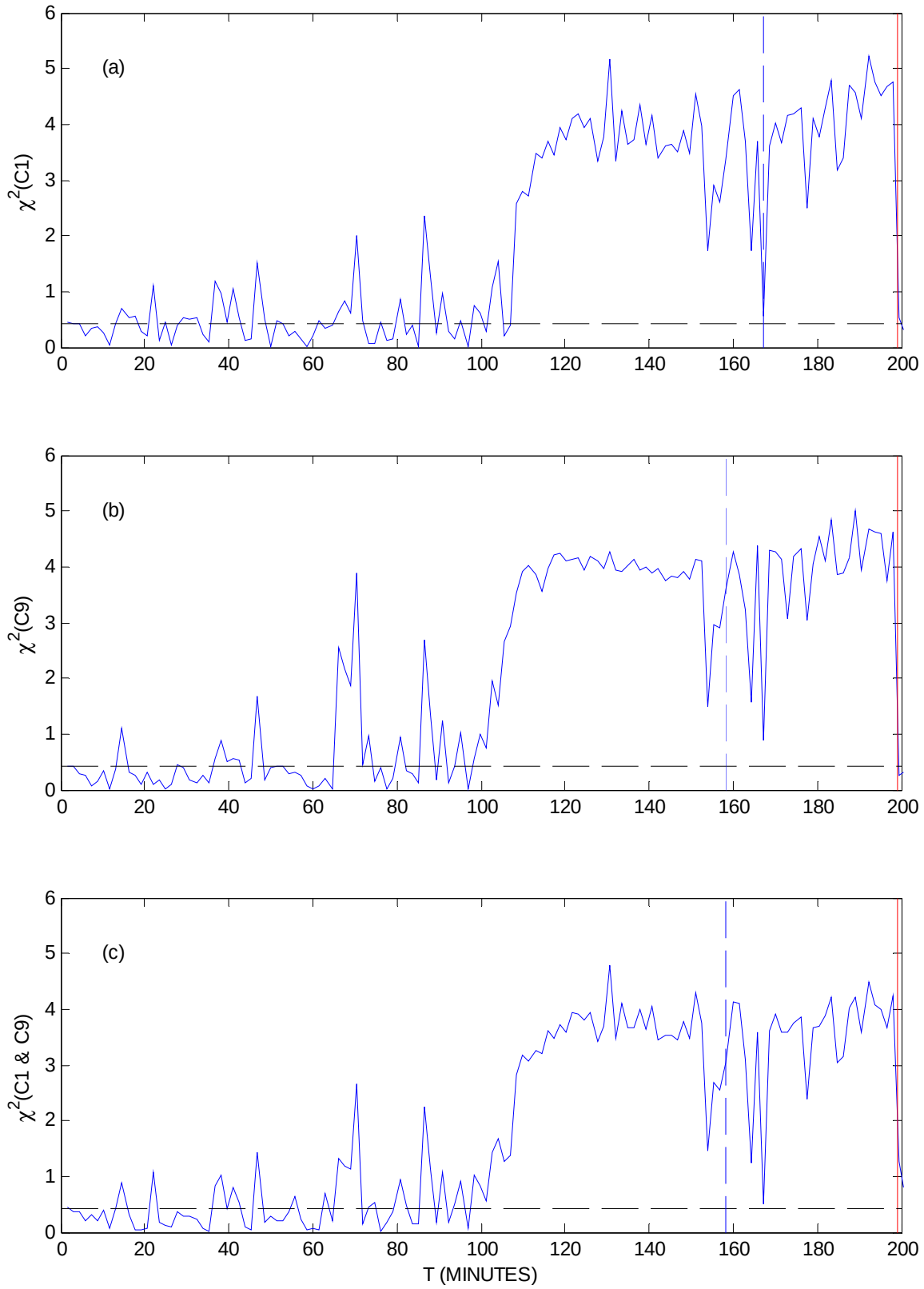

Figure 3: Typical values of the χ^2 dissimilarity measure for dataset #264 in (a) channel C1 only, (b) channel C9 only, and (c) the dual-channel analysis of C1 and C9. The seizure onset occurs at 199 minutes (solid vertical line). The forewarning threshold is U_C=0.419 (N_{OCC}=42), as denoted by the horizontal dashed lines. The corresponding forewarning times are 31.8, 40.6, and 40.6 minutes before the seizure event in subplots (a)-(c), respectively. Thus, the dual-channel analysis extracts the better forewarning from C9 rather than the shorter one from C1.

Figure 1. Layout of the EEG scalp electrodes in the standard 10-20 montage, looking at the patient's head from above. The bipolar montage uses signal differences from adjacent electrodes, such as $C1 = F_7 - FP_1$ and $C_9 = F_8 - FP_2$, which are the symmetric channels, which are chosen for the multi-channel analysis via Eq. (2).

Figure 2. Distribution of forewarning time (T_{FW}) for the dual-channel analysis of bipolar scalp EEG with f_T = 56/60. The corresponding sensitivity, TP/(TP + FN), is 39/40; the specificity, TN/(TN + FP), is 17/20. The stair-step (solid) plot indicates the relative occurrence frequency (arbitrary units) of T_{FW} versus time intervals in one-hour increments, beginning at 0-1 hour, and ending with 5-6 hours. The monotonically increasing (dashed) curve is the cumulative distribution function (CDF) of T_{FW} versus time. The H-bar with the star in the middle indicates the mean value of T_{FW} (2.3 hours) and the standard deviation of the mean.

Figure 3. Typical values of the χ^2 dissimilarity measure for dataset #264 in (a) channel C1 only, (b) channel C9 only, and (c) the dual-channel analysis of C1 and C9. The seizure onset occurs at 199 minutes (solid vertical line). The forewarning threshold is U_C=0.419 (N_{OCC}=42), as denoted by the horizontal dashed lines. The corresponding forewarning times are 31.8, 40.6, and 40.6 minutes before the seizure event in subplots (a)-(c), respectively. Thus, the dual-channel analysis extracts the better forewarning from C9 rather than the shorter one from C1.