International Journal of Bioelectromagnetism
Vol. 9, No. 3, pp. 134 - 145, 2007
www.ijbem.org

Assessment of Measures of Scalar Time Series Analysis in Discriminating Preictal States

Dimitris Kugiumtzis[a], Ioannis Vlachos[a], Aggeliki Papana[a], Pål G Larsson[b]
[a]*Department of Mathematical, Physical and Computational Sciences, Faculty of Technology, Aristotle University of Thessaloniki, Thessaloniki, Greece*
[b]*Department of Neurodiagnostics, Division of Clinical Neuroscience, Rikshospitalet University Hospital, Oslo, Norway*

Correspondence: D Kugiumtzis, Department of Mathematical, Physical and Computational Sciences, Faculty of Technology, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece. E-mail: dkugiu@gen.auth.gr, phone +30 2310 995955, fax +30 2310 995958

Abstract. We evaluate a large number of measures used in the statistical and nonlinear analysis of EEG as to their power in discriminating different preictal states. Some new measures are proposed and the analysis involves also feature time series, such as local minima and maxima, the time between them, as well as interspike intervals. The measures are computed on consecutive non-overlapping segments of 30 s of multichannel scalp EEG at the preictal period of several hours. The discrimination power of each measure is assessed by the receiver operating characteristic (ROC) curve computed on measure values grouped in different preictal stages. The results on 7 epileptic EEG records show that simple measures have the same, and at cases better, power in discriminating preictal states than other more sophisticated measures.

Keywords: Preictal EEG; feature time series; nonlinear analysis; correlation; complexity; receiver operating characteristic

1. Introduction

The prediction of epileptic seizure from electroencephalograph recordings (EEG) is an open issue and there is still no method that can reliably provide in a reliable way such predictions. The advance of nonlinear dynamics and chaos, along with other fields, such as synchronization, networks, random matrix theory, and statistical testing, have given rise to measures and approaches that have been used in the analysis of preictal EEG with varying success [Andrzejak et al., 2003; Stam, 2005; Iasemidis et al., 2005]. Over the last years the propositions of measures for seizure prediction have escalated but only few works focused on evaluating and comparing even a subset of these measures, as in [Mormann et al., 2005], leaving a gap of validation of the proposed measures. Very recently, critical reviews have pointed that there is a long way to go before we can establish statistical significance for the predictability of any measure in terms of its sensitivity and specificity [Mormann et al., 2006; Mormann et al., 2007; Gudmundsson et al., 2007]. These remarks are supported by the failure of any approach to score adequately in the recent competition of seizure prediction (Seizure Prediction Competition, Freiburg, October 2007).

There is an obvious need to review the large number of proposed measures before any of them can be used in clinical testing. Along these lines, recently a number of different measures of scalar time series analysis were tested in discriminating early and late preictal stages in [Kugiumtzis et al., 2006]. We extend this study to include most of the measures that have been used so far in univariate EEG analysis. Moreover, we modify some of the known measures of correlation and propose some new measures as well. On top of these measures that are estimated directly on the oscillating EEG time series, we extract from the original EEG 7 feature time series that regard different characteristics of the oscillating EEG signal, such as the local minima and maxima. We compute on the feature time series some simple statistics as well as some of the same measures computed on the original EEG. An indisputable advantage of the measures on the feature time series is that are computationally efficient as the length of the feature time series is a small fraction of the original EEG.

A complete and exhaustive evaluation of all these measures would include a good number of different EEG measurements (scalp, intracranial), epilepsy seizure types, physiological conditions, data pre-processing, and varying duration of the tested preictal period, and all these require an excessive amount of resources and computation time. Here the evaluation study is confined to 7 scalp multi-channel EEG recordings of preictal periods lasting at least 3 h and all seizures are generalized tonic clonic. The interest is in the power of the measures in discriminating late preictal state (just before seizure onset) to earlier preictal states. For this the statistic of the area under the receiver operating characteristic (ROC) curve is used and the measures are ranked according to their ROC-relate scores.

The structure of the paper is as follows. In Section 2, the data and the measures are presented and in Section 3 the results of the evaluation of the measures with regard to score ranking are given. The results are discussed and conclusions are drawn in Section 4.

2. Material and Methods

2.1. The EEG data

The EEG data were recorded on a routine basis at the Department of Neurodiagnostics, Rikshospitalet University Hospital, Norway. They comprise of 7 preictal multi-channel scalp EEG records, each regarding a single episode of a generalized tonic clonic seizure from a different patient, covering at least 3 h prior to the seizure onset. Recordings with 63 channels, according to the 10-10 system are used for the two first EEG sets (denoted A and B), while for the other 5 EEG sets (denoted from C to G) recordings were done with 25 channels according to the 10–20 system with added low rows. Reference electrode was placed at CP1 as defined in 10 – 10. Fronto-polar channels are excluded because they tend to contain continuous artefacts, bringing down the number of channels to 53 and 19, respectively. As all methods considered in this study do not rely on very high frequency components, the data were high-pass filtered at 0.3 Hz, low-pass filtered at 40 Hz, and down-sampled at 100 Hz (0.01 s). The latter was done to improve computation efficiency. No procedure for artefact removal or any other form of pre-processing was applied in order to assess the methods under standard conditions of clinical practice.

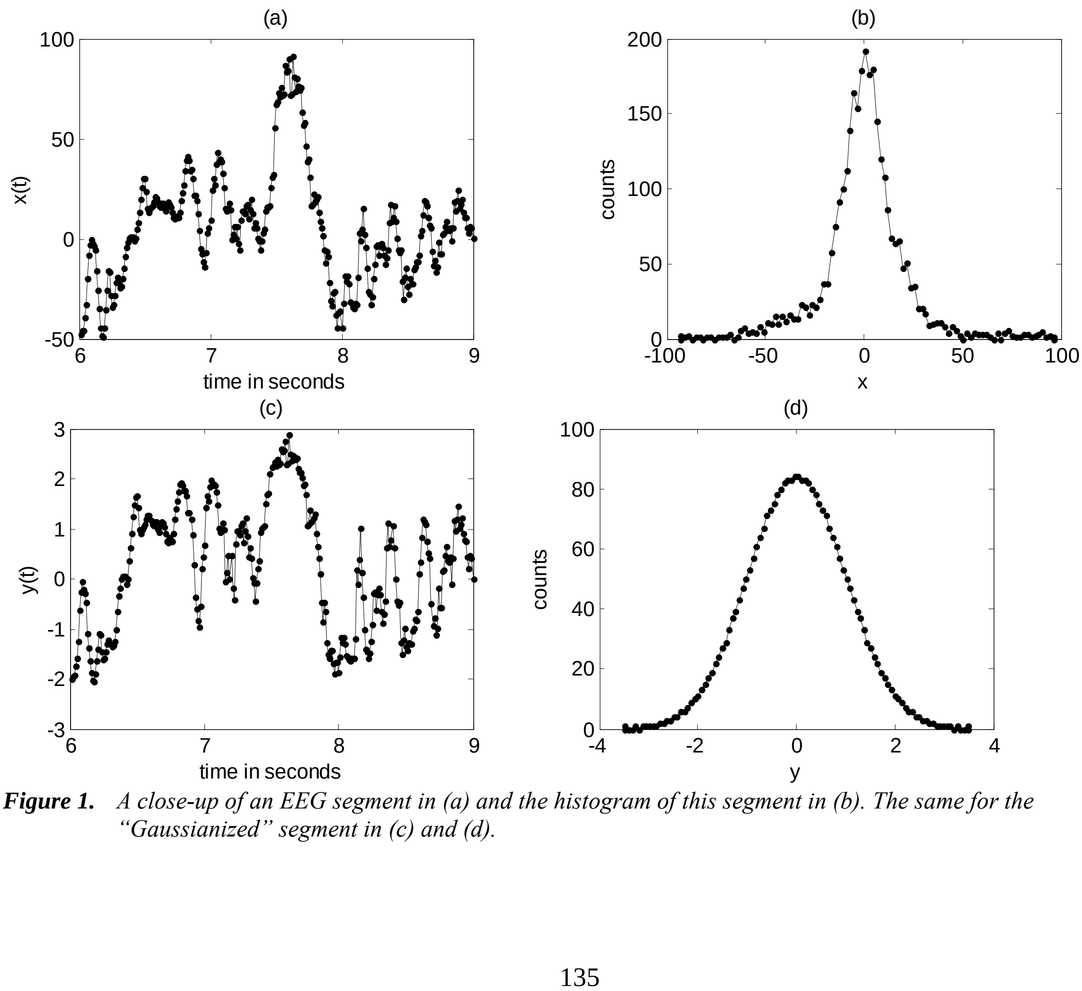

Figure 1. *A close-up of an EEG segment in (a) and the histogram of this segment in (b). The same for the "Gaussianized" segment in (c) and (d).*

Each record was split into consecutive non-overlapping segments of duration 30 s giving N=3000 samples at each segment, denoted $\{x_t\}_{t=1}^N$. The EEG signal exhibits irregular fluctuations and has seldom Gaussian amplitude, i.e. the empirical marginal cumulative density function (cdf) $F_x(x_t)$ of a segment $\{x_t\}_{t=1}^N$ deviates from Gaussian. This can be seen from the histogram of a typical $\{x_t\}_{t=1}^N$ in Fig.1b (a close-up of the segment is shown in Fig.1a). A large deviation from Gaussian marginal distribution can affect measures sensitive to the amplitude distribution. We therefore consider a static transform of $\{x_t\}_{t=1}^N$ to the "Gaussianized" time series $\{y_t\}_{t=1}^N$ that possesses Gaussian marginal cdf $\Phi(y_t)$, and it is given as $y_t = \Phi^{-1}\left(F_x(x_t)\right)$. As shown in Fig.1c and d, this transform renders the amplitude normal (Gaussian) without alerting the dynamics of the time series. It was observed in [Kugiumtzis et al., 2006] that correlation measures would perform discrimination tasks on preictal EEG better when the data were Gaussianized. Therefore many of the measures were computed on $\{y_t\}_{t=1}^N$ instead of $\{x_t\}_{t=1}^N$.

2.2. Measures on the original EEG

A large number of measures are computed on the oscillating EEG time series. The correlation measures are estimated on $\{y_t\}_{t=1}^N$ to eliminate the effect of heavy tails and outliers in the signal amplitude distribution. A list of the correlation measures is given in Table 1.

Table 1. *Correlation measures on the Gaussianized EEG time series. Each line regards a measure type and the number of measures derived for different parameter values is given in the third column.*

Symbol	*Description*	#
τ_e	decorrelation time, lag for $r_y^{\mathrm{P}}(\tau_e)=1/e$ (see below)	1
$\tau_{\max}$	lag of levelling of mutual information (see below)	1
$r_y^{\mathrm{P}}(\tau)$ $Cr_y^{\mathrm{P}}(40), Cr_y^{\mathrm{P}}(\tau_{\max})$	(Pearson) autocorrelation at $\tau = 5,10,20,30$ Cumulative autocorrelation for lag 40 and $\tau_{\max}$	4 2
$r_y^{\mathrm{S}}, Cr_y^{\mathrm{S}}$	Same as above for Spearman autocorrelation	6
$r_y^{\mathrm{K}}, Cr_y^{\mathrm{K}}$	Same as above for Kendall autocorrelation	6
I_y, CI_y	Same as above for mutual information	6
dI_y, dCI_y	Same as above for $I_y - I_y^g$ where $I_y^g = -0.5\ln(1-r_y^{\mathrm{P}})$	6
$r_y^{\mathrm{b}}, Cr_y^{\mathrm{b}}$	Same as above for bicorrelation (not for $\tau = 30$)	5

The standard correlation coefficients of Pearson, Spearman and Kendall are used as measures of autocorrelation at different lags ($\tau = 5,10,20,30$) and the decorrelation time, i.e. the lag at which the Pearson autocorrelation function first equals to 1/e, is also included as a measure of memory in the system [Martinerie et al., 1998; Ehlers et al., 1998; Andrzejak et al., 2006]. In addition, the cumulative autocorrelation is computed (for Pearson, Spearman and Kendall), summing the square of autocorrelation up to a maximum lag. Two choices of maximum lag are considered, a fixed value (set to 40) and the lag for which the mutual information is leveling off according to a criterion for flattening. The latter is also used as a measure. Mutual information and cumulative mutual information are used in order to quantify both linear and nonlinear dependencies for the same lags as above, where the estimation is based on histogram of 16 equidistant bins [Chillemi et al., 2003; Nicolaou and Nasuto, 2005; Papana and Kugiumtzis, 2007]. In an attempt to include a measure of solely nonlinear correlations we compute the difference between the mutual information at some lag from the normalized time series $\{y_t\}_{t=1}^N$, I_y, and the "Gaussian" mutual information I_y^g, i.e. the mutual information as if $\{y_t\}_{t=1}^N$ is from a Gaussian process with autocorrelation r given as $I_y^g = -0.5\ln(1-r)$ [Darbellay and Vajda, 1999]. We include also a different nonlinear autocorrelation measure based on the third-order joint moment, referred to as three-point autocorrelation or bicorrelation and compute it at lags $\tau = 5,10,20$. Moreover, we consider the cumulative bicorrelation by summing the bicorrelations up to lag 40.

Another class of measures listed in Table 2 regard other characteristics of the underlying system, such as dimension, entropy and complexity. Some of these measures are based on the nonlinear dynamics theory and require the reconstruction of state space from the scalar time series with the method of delays determined by the embedding dimension m and the delay τ [Kantz and Schreiber, 1997]. A popular measure of fractal dimension is the correlation dimension, but reliable estimation of it requires scaling of the correlation integral, $C(r)$ with the distance r that hardly can be maintained on preictal EEG data. Therefore we consider as a measure of point density the $C(r)$ at a fixed distance (r=0.2), in line with other studies on EEG [Lerner et al., 1996; McSharry et al. 2003; Andrzejak et al., 2006]. Our modification in using $\{y_t\}_{t=1}^{N}$ instead of the original EEG in the computation of $C(r)$ has also an operational advantage: it suppresses the dependence of r on the amplitude changes in the original EEG. Moreover, we use a reconstructed state space of low and high dimension (m=5,10) for a fixed delay ($\tau = 10$). In an inverse manner, we consider the distance r that regards $C(r)$=0.1, which has been used as another measure of determinism [Ehlers et al., 1998]. A different measure based on inter-point distance and neighborhood in the reconstructed state space is the measure of false nearest neighbors (FNN) that estimates the percentage of false nearest neighbors at each embedding dimension by tracking their distance at the augmented state space increased by one dimension [Kennel et al., 1992]. We compute this percentage at m=5,10. A popular measure of chaos is the estimation of the largest Lyapunov exponent λ, which has been one of the first measures based on nonlinear dynamics to be applied for the prediction of epileptic seizures with reporting success until to-date [Iasemidis et al., 1991; Iasemidis et al., 2005; patent Sackellares et al., 2007]. Here we apply a standard computation scheme for λ from the Lyapunov spectrum estimation as implemented in the non-commercial software package TISEAN [Hegger et al., 1999], using two neighborhood sizes (20 and 40 neighbors) and state space reconstruction parameters as for the other measures.

Table 2. *Measures of dimension, entropy and complexity listed as in Table 1.*

Symbol	*Description*	#
$C_y(r,m)$	correlation integral for r=0.2 and m=5,10	2
$r(C_y,m)$	distance for $C_y(m)=0.1$ for m=5,10	2
$\mathrm{FNN}_x(m)$	% of false nearest neighbors, τ =10 and m=5,10	2
$\lambda_y(m,K)$	Largest Lyapunov exponent using local maps τ =10, m=5,10 and K=20,40	4
$\mathrm{AE}_y(r,m)$	Approximate entropy for r=0.2 and m=5,10	2
$\mathrm{AC}_x^{\mathrm{d}}$ $\mathrm{AC}_x^{\mathrm{p}}$ $\mathrm{AC}_y^{\mathrm{d}}$	Algorithmic complexity: 16 equidistant and equiprobable bins. The same for "Gaussianized" data and 16 equidistant bins	2 1
HjM_y, HjC_y	Hjorth parameters: mobility and complexity	2
SE_y	Spectral entropy (frequencies in [0.5,48])	1
$\mathrm{E}_y(i)$ i=$\delta,\theta,\alpha,\beta,\gamma$	Energy in $\delta,\theta,\alpha,\beta,\gamma$ frequency bands	5
mf	median frequency (frequencies in [0.5,48])	1
HeS_y, HeE_y, HeC_y	Hurst exponent: Slope, Fit error, Relative Change	3
DFS_y, DFE_y, DFC_y	Detrended Fluctuation Analysis: as above	3

In the same list of measures, the approximate entropy is a standard measure of entropy and system complexity [Burioka et al., 2003], and we compute it here for the same parameters as for $C(r)$. Other measures of complexity included in our study are the algorithmic complexity and the so-called Hjorth mobility and complexity parameters. The algorithmic complexity measures the complexity of a sequence of symbols derived from the original data by partitioning its values in bins and assigning a symbol at each bin [Ziv and Lempel, 1978]. We use 16 bins and both equidistant and equiprobable partitioning. The Hjorth parameters are simple measures of signal complexity given by the second (mobility) and fourth (complexity) statistical moments of the power spectrum and quoting from [Hjorth, 1970] they are 'clinically useful tools for the quantitative description of an EEG'.

Measures calculated on the frequency domain are:

1. the spectral entropy, i.e. the entropy computed on the power spectrum of the EEG data [Rautte et al., 2005].
2. the spectral band power on frequency bands corresponding to brain waves delta, theta, alpha, beta and gamma as defined in classical EEG analysis
3. the frequency up to which 50% of the total spectral power is accumulated, referred to as median frequency.

Finally, we include in this list long range correlation measures that are used in the analysis of non-stationary time series, mainly in finance and climate. The Hurst exponent [Dangel et al., 1999; Nurujjaman et al., 2007] and the Detrended Fluctuation Analysis (DFA) [Peng et al., 1995] are measures for estimating the extent of long correlations in processes and characterize non-stationary time series with regard to their self-similarity. For each of the two approaches we derive three measures from the graph of a variance-like quantity vs the time window length, the slope estimate, the error of the linear fit that gives the slope estimate, and the cumulative relative change of the fit error when increasing the window for fitting.

Table 3. *Measures based on linear and nonlinear modelling listed as in Table 1.*

Symbol	*Description*	#
$\mathrm{dAR}_x(m_1,m_2)$	Difference in MSE of AR at orders *m*=0,5,10	2
$\mathrm{dloc}_x(m_1,m_2,K)$	Same as above for local-AR and *K*=20,40	4
$\mathrm{dlocAR}_x(m,K)$	Difference local-AR and AR, *m*=5,10, *K*=20,40	4

The third class of measures given in Table 3 contains measures derived from modelling. Linear autoregressive models (AR) and local-linear models (local-AR) of varying orders were fitted to the original EEG time series and the relative change of goodness-of-fit between them was computed. The goodness-of-fit criterion was the mean squared error (MSE) and the selected orders were *m*=0, 5, 10 with *K*=20 and 40 neighbours for the local linear models. The measures are given by the relative change of MSE when going from one type of model to another (either from linear to non-linear of same order or from small order to higher order) [Nielsen and Madsen, 2001]. Note that the use of *m*=0 here is in order to facilitate the presentation of measures, i.e. the change of MSE for AR from order 0 to 5 is actually the MSE for AR of order 5.

2.3. Measures on the feature time series

Basic analysis tools that account for the oscillations in the EEG signal are the Fourier, wavelet and, recently, empirical mode decomposition. An alternative new approach along these lines that we first introduced in [Kugiumtzis et al., 2006] is to extract directly features from each oscillation of the EEG signal and form feature time series. Seven time series of features are extracted from $\{x_t\}_{t=1}^{N}$ that are listed and denoted in Table 4.

Table 4. *Feature time series extracted from the EEG oscillating signal* $\{x_t\}_{t=1}^{N}$.

Time series	*Description*
$\left\{x_i^{\max}\right\}_{i=1}^{n}$	local maxima (from window of 15 samples)
$\left\{x_i^{\min}\right\}_{i=1}^{n}$	local minima
$\left\{t_i^{\mathrm{osc}}\right\}_{i=1}^{n}$	time between consecutive local maxima
$\left\{t_i^{\max,\min}\right\}_{i=1}^{n}$	time from local maximum to next local minimum
$\left\{\delta_i^{\max,\min}\right\}_{i=1}^{n}$	difference of local maxima and minima
$\left\{z_i^{\mathrm{abs}}\right\}_{i=1}^{n'}$	inter-spike interval on $\left\{\mid x_t \mid\right\}_{t=1}^{N}$
$\left\{z_i^{\mathrm{cut}}\right\}_{i=1}^{n''}$	inter-spike interval on $\left\{x_t + \mathrm{cutoff}\right\}_{t=1}^{N}$

Regarding the turning points (local maxima and minima), a sample of the EEG time series is identified as local maximum (minimum) if it is the largest (smallest) of the samples in a window of 15 samples centered on this sample. An earlier study in [Kugiumtzis et al., 2006] showed that smaller windows gave similar results. The four first features in Table 4 determine the basic form of an oscillation in terms of magnitude and time though in EEG the lines connecting the turning points may not be smooth and contain glitches. The fifth feature of the difference of successive turning points is included in order to register information about the range of the oscillation and to account for small drifts in the oscillating time series. The inter-spike intervals (defined in two ways to treat negative values in the last two lines of Table 4) characterize the oscillations in a different way. Note that the name does not refer directly to the epileptic spikes in EEG. An inter-spike interval (ISI) is defined as the time needed for the sum of the time series values to reach a given threshold [Sauer, 1994]. A series of consecutive inter-spike interval times is obtained with a threshold determined by the mean oscillation time of the time series and the mean of its values.

The measures estimated on the seven time series are listed in Table 5, grouped as for the measures on the oscillating EEG time series. We include also a group of simple statistics of all 7 feature time series, namely the median and inter-quartile range (these were found to be more robust measures of central tendency and variability than the mean and variance or standard deviation, respectively). We select only few of the measures on the oscillating time series that we expect to have robust estimates for due to the severely reduced length of the feature time series. For the same reason, smaller parameter values are used. As correlation measures, we use only the Pearson autocorrelation but include the cross-correlation as well as the cumulative cross-correlation of the four features that capture the basic form of the oscillation. We use the algorithmic complexity as a complexity measure, the approximate entropy as an entropy measure and the FNN as a dimension measure. The modeling-based measures are as for the oscillating time series but for smaller embedding dimension *m* and number of neighbors *K*. Including the different parameter values the total number of measures on oscillating and feature time series is 284.

Table 5. *Measures on the seven feature time series, where w denotes any of them. The number of measures derived for different parameter values and on different feature time series is explained in the third column.*

Symbol	*Description*	#
	Simple Statistics	
$\mathrm{m}(w)$	median of $\{w_i\}_{i=1}^n$	1x7
$\mathrm{IQR}(w)$	inter-quartile range of $\{w_i\}_{i=1}^n$	1x7
	Correlation	
$r_w^{\mathrm{P}}(\tau)$	(Pearson) autocorrelation for $\tau = 1,2,3$	3x7
$Cr_w^{\mathrm{P}}(3)$	Cumulative autocorrelation for lag 3	1x7
$c_{w_1,w_2}(\tau)$	Cross-correlation for τ =0,1,2,3, only for $x_i^{\max}, x_i^{\min}, t_i^{\mathrm{osc}}, t_i^{\max,\min}$	4x3x4
$Cc_{w_1,w_2}(3)$	Cumulative cross-correlation for lag 3, only for $x_i^{\max}, x_i^{\min}, t_i^{\mathrm{osc}}, t_i^{\max,\min}$	3x4
	Complexity / Entropy / Dimension	
$\mathrm{AC}_w^{\mathrm{p}}$	Algorithmic complexity equiprobable binning using 2 bins	1x7
$\mathrm{AE}_w(r,m)$	Approximate entropy for *r*=0.2 and *m*=1,2	2x7
$\mathrm{FNN}_w(m)$	Percentage of false nearest neighbors, *m*=1,2	2x7
	Modeling	
$\mathrm{dAR}_w(m_1,m_2)$	Difference in MSE of AR, *m*=0,1,2	2x7
$\mathrm{dloc}_w(m_1,m_2,K)$	Same as above for local-AR and *K*=5,10	4x7
$\mathrm{dlocAR}_w(m,K)$	Difference local-AR - AR, *m*=1,2 and *K*=5,10	4x7

2.4. Evaluation of measures

The estimation of each measure on all consecutive segments over a period of at least 3 h gives the measure profile (a series of at least 360 values for each measure). An example of the profile of a simple and more involved measure (the median of the local maxima $\mathrm{m}\left(x_i^{\max}\right)$ and the largest Lyapunov exponent $\lambda_y(5,20)$, respectively) estimated on the EEG signal from a temporal and a central channel is shown in Fig. 2a and b. It appears that better discrimination of late preictal state from earlier preictal states can be established with the simple measure of $\mathrm{m}\left(x_i^{\max}\right)$.

In [Kugiumtzis et al. 2006] the discriminating power of some of the listed measures was assessed by means of statistical comparison of two groups (t-test, nonparametric tests) regarding smaller preictal periods. Here, we consider larger periods that span 30 min and resulting in 60 samples in each group giving a sufficient estimate of the distribution of the measure for this period (preictal state). Thus the comparison can extent to the whole distribution rather than the mean or median, as shown in Fig. 2c, and we consider for this the receiver operating characteristic curve (ROC) [Armitage et al., 2002]. The ROC curve gives the level of overlapping or distance of two sample distributions, so that the closer to the diagonal the ROC curve lies the more the two distributions overlap (see Fig. 2d). To quantify the level of overlapping we use here the area under the ROC curve (AUC), so that AUC values close to one suggest disjoint distributions, and for our case clear discrimination of the two preictal states, with regard to the measure under study. We verified that relative entropy [Cover and Thomas, 1991] quantifies similarly with AUC the level of overlapping (distance) of the two distributions and we therefore used only AUC in the evaluation procedure.

We note that ROC analysis requires that the samples in each group are independent. However, there may be the case that significant correlations within a preictal state are found, probably due to trends within the period, such as the upward trend in the late preictal state in Fig. 2a. Since such a trend is itself evidence in favour of discrimination, we overlook the independence assumption in the computation of ROC and AUC.

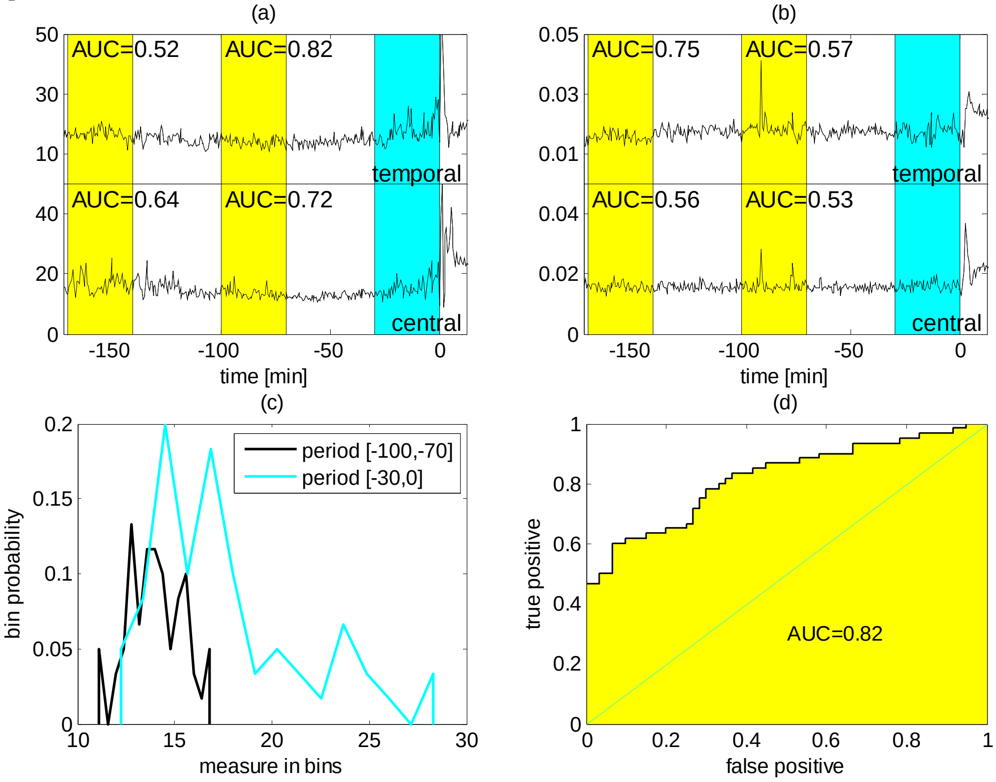

Figure 2. *Profiles of the median of local maxima in (a) and the largest Lyapunov exponent in (b) for a temporal and a central channel (upper and lower panel). The AUC values displayed at each panel regard the comparison of the periods in first and second light shaded areas to the third period of shaded area ending at seizure onset. The histogram of the samples in the second and third period in the upper panel of (a) are superimposed in (c) and the ROC curve and AUC area for these two periods are drawn in (d).*

3. Results

We consider groups of measure values at different periods of 30 min long (60 values) that regard different preictal stages. As the interest is in discriminating the late preictal state from earlier preictal states, we compare by means of AUC the group of the late preictal state in the period [-30,0] (with reference to the seizure onset at time 0 as identified by a clinician), with other groups at earlier periods of the same length. For each group comparison, the AUC of a measure is computed for every channel and the average AUC is the score of the measure.

The 10 measures scoring highest in terms of the average AUC for the comparison of groups in the periods [-30,0] and [-190,-160] are listed for all 7 seizure episodes in Table 6. It can be seen that there is large variability of the AUC scores across measures as well as across seizure episodes. While it is somehow expected that no measure singles out, it appears that some measures occur often among the highest ranked measures for a range of seizure episodes, such as the autocorrelation at lag 10 (Spearman, Kendall and Pearson) and the median of inter-spike intervals. Interestingly, more advanced methods, such as chaos-related measures and modeling-based measures are not included among the best. The high AUC values of a number of measures would suggest that discrimination is possible between these two preictal stages in all of the 7 seizure episodes. However, among 284 measures it is plausible that some measures score high in AUC. To investigate this further, we repeated the same

AUC calculations using the same late preictal state and other earlier preictal states and found that there was some variation in the AUC scores of the measures but the same simple measures performed similarly.

We computed the grand average of AUC from comparisons of the late preictal stage to 5 consecutive stages of 30 min long, starting at 190 min prior to seizure onset and ending at 40 min prior to seizure onset. The results from the ranking of the measures are shown in Table 7. Though the ranking of the measures is not the same as for a single comparison (compare Table 6 to Table 7) we note that the same simple measures occur consistently among the 10 best. Again the highest AUC scores suggest significant discrimination of the late preictal state from the earlier preictal record.

Table 6. *The 10 measures ranked highest in terms of their AUC values from period [-30,0] and [-190,-160], where seizure onset is at time 0. For each of the 7 seizure episodes, the measure is given together with its AUC value.*

	A		**B**		**C**		**D**	
rank	*measure*	*AUC*	*measure*	*AUC*	*measure*	*AUC*	*measure*	*AUC*
1	$\mathrm{m}(z_i^{\mathrm{abs}})$	0.743	HjC_y	0.803	$\mathrm{m}(x_i^{\max})$	0.827	$\mathrm{m}(z_i^{\mathrm{abs}})$	0.919
2	$r_y^{\mathrm{K}}(10)$	0.735	$\mathrm{E}_y(\beta)$	0.757	$\mathrm{m}(x_i^{\min})$	0.815	$\mathrm{m}(x_i^{\max})$	0.890
3	$r_y^{\mathrm{S}}(10)$	0.734	$\mathrm{dlocAR}_x(5,20)$	0.755	$\mathrm{E}_y(\gamma)$	0.810	$\mathrm{m}(x_i^{\min})$	0.889
4	$\mathrm{IQR}(x_i^{\max})$	0.721	$\mathrm{dlocAR}_x(10,20)$	0.750	HjM_y	0.801	$\mathrm{m}(\delta_i^{\max,\min})$	0.889
5	$r_y^{\mathrm{P}}(10)$	0.720	$\mathrm{E}_y(\gamma)$	0.749	$\mathrm{m}(\delta_i^{\max,\min})$	0.726	$\mathrm{IQR}(x_i^{\min})$	0.856
6	$\mathrm{IQR}(x_i^{\min})$	0.719	$\mathrm{dlocAR}_x(5,40)$	0.747	$\mathrm{m}(z_i^{\mathrm{abs}})$	0.715	$\mathrm{IQR}(x_i^{\max})$	0.853
7	$r_y^{\mathrm{S}}(5)$	0.704	$\mathrm{dlocAR}_x(10,40)$	0.737	$\mathrm{AC}_y^{\mathrm{d}}$	0.699	$\mathrm{IQR}(z_i^{\mathrm{abs}})$	0.845
8	$r_y^{\mathrm{K}}(5)$	0.703	$\mathrm{E}_y(\theta)$	0.711	$\mathrm{AC}_x^{\mathrm{p}}$	0.697	$\mathrm{IQR}(\delta_i^{\max,\min})$	0.833
9	$r_y^{\mathrm{P}}(5)$	0.694	$r_y^{\mathrm{S}}(5)$	0.705	$\mathrm{dloc}_x(5,10,20)$	0.696	$\mathrm{dAR}_x(0,5)$	0.812
10	$I_y(5)$	0.693	$r_y^{\mathrm{K}}(5)$	0.704	$\mathrm{m}f$	0.694	$\mathrm{dlocAR}_x(10,40)$	0.781

	E		**F**		**G**	
rank	*measure*	*AUC*	*measure*	*AUC*	*measure*	*AUC*
1	DFC_y	0.850	τ_e	0.779	τ_e	0.972
2	$r_y^{\mathrm{S}}(10)$	0.847	$r_y^{\mathrm{S}}(10)$	0.677	HjC_y	0.920
3	$r_y^{\mathrm{K}}(10)$	0.846	$r_y^{\mathrm{K}}(10)$	0.675	$r_y^{\mathrm{K}}(30)$	0.914
4	$r_y^{\mathrm{P}}(10)$	0.843	$r_y^{\mathrm{P}}(10)$	0.652	$r_y^{\mathrm{S}}(30)$	0.913
5	HeC_y	0.816	$I_y(10)$	0.646	$r_y^{\mathrm{S}}(5)$	0.913
6	$\mathrm{E}_y(\theta)$	0.815	$dI_y(10)$	0.645	$r_y^{\mathrm{K}}(5)$	0.913
7	$r_y^{\mathrm{S}}(20)$	0.802	$Cr_y^{\mathrm{S}}(40)$	0.636	$r_y^{\mathrm{P}}(30)$	0.911
8	$r_y^{\mathrm{K}}(20)$	0.802	$\mathrm{IQR}(t_i^{\max,\min})$	0.633	$r_y^{\mathrm{P}}(5)$	0.899
9	$r_y^{\mathrm{P}}(20)$	0.794	$Cr_y^{\mathrm{K}}(40)$	0.633	$\mathrm{E}_y(\delta)$	0.886
10	$\mathrm{E}_y(\alpha)$	0.778	$Cr_y^{\mathrm{S}}(\tau_{\max})$	0.629	$\mathrm{m}f$	0.884

Table 7. *As in Table 6 but for the grand average of AUC from 5 comparisons of the period [-30,0] to periods [-190,-160], [-160,-130], [-130,-100], [-100,-70] and [-70,-40].*

	A		**B**		**C**		**D**	
rank	*measure*	*AUC*	*measure*	*AUC*	*measure*	*AUC*	*measure*	*AUC*
1	$\mathrm{m}(z_i^{\mathrm{abs}})$	0.775	$\mathrm{E}_y(\beta)$	0.683	$\mathrm{E}_y(\gamma)$	0.774	τ_e	0.822
2	$\mathrm{IQR}(x_i^{\min})$	0.739	HjC_y	0.676	$\mathrm{dlocAR}_x(10,20)$	0.742	$\mathrm{m}(z_i^{\mathrm{abs}})$	0.764
3	$\mathrm{IQR}(x_i^{\max})$	0.735	$\mathrm{E}_y(\gamma)$	0.666	HjM_y	0.715	$\mathrm{m}(x_i^{\min})$	0.734
4	$\mathrm{m}(\delta_i^{\max,\min})$	0.735	$\mathrm{dlocAR}_x(10,40)$	0.663	$\mathrm{m}f$	0.704	$\mathrm{m}(\delta_i^{\max,\min})$	0.734
5	$\mathrm{IQR}(z_i^{\mathrm{abs}})$	0.732	$\lambda_y(5,40)$	0.663	$\mathrm{dlocAR}_x(5,20)$	0.700	$\mathrm{m}(x_i^{\max})$	0.732
6	$\mathrm{m}(x_i^{\max})$	0.731	$\mathrm{dlocAR}_x(5,40)$	0.657	$\mathrm{m}(x_i^{\max})$	0.693	$\mathrm{IQR}(z_i^{\mathrm{abs}})$	0.714
7	$\mathrm{m}(x_i^{\min})$	0.719	$\mathrm{AC}_x^{\mathrm{p}}$	0.652	$\mathrm{m}(x_i^{\min})$	0.683	$\mathrm{IQR}(x_i^{\max})$	0.713
8	$\mathrm{IQR}(\delta_i^{\max,\min})$	0.692	$\mathrm{AC}_y^{\mathrm{d}}$	0.649	$\mathrm{dloc}_x(5,10,20)$	0.677	$\mathrm{IQR}(x_i^{\min})$	0.711
9	$\mathrm{AE}_{z_{\mathrm{cut}}}(r,m)$	0.673	$\mathrm{dlocAR}_x(5,20)$	0.646	$\mathrm{E}_y(\beta)$	0.668	$\mathrm{dAR}_x(0,5)$	0.700
10	$r_y^{\mathrm{K}}(10)$	0.661	$\mathrm{dlocAR}_x(10,20)$	0.645	$\mathrm{E}_y(\delta)$	0.661	$r_y^{\mathrm{K}}(5)$	0.698

	E		**F**		**G**	
rank	*measure*	*AUC*	*measure*	*AUC*	*measure*	*AUC*
1	DFC_y	0.873	τ_e	0.786	HjC_y	0.939
2	$r_y^{\mathrm{K}}(20)$	0.864	$\mathrm{E}_y(\gamma)$	0.652	$r_y^{\mathrm{K}}(30)$	0.914
3	$r_y^{\mathrm{S}}(20)$	0.864	$\mathrm{m}(x_i^{\min})$	0.65	$r_y^{\mathrm{S}}(30)$	0.914
4	$r_y^{\mathrm{P}}(20)$	0.855	$\mathrm{m}(\delta_i^{\max,\min})$	0.647	$r_y^{\mathrm{P}}(30)$	0.908
5	HeC_y	0.829	$Cr_y^{\mathrm{S}}(40)$	0.644	$r_y^{\mathrm{S}}(5)$	0.904
6	$\mathrm{AC}_y^{\mathrm{d}}$	0.814	$c_{x^{\max},x^{\min}}(1)$	0.644	$r_y^{\mathrm{K}}(5)$	0.904
7	$\mathrm{AC}_x^{\mathrm{p}}$	0.814	$Cr_y^{\mathrm{K}}(40)$	0.642	$r_y^{\mathrm{P}}(5)$	0.891
8	$\mathrm{E}_y(\theta)$	0.814	$dCI_y(40)$	0.641	$\mathrm{E}_y(\alpha)$	0.874
9	$r_y^{\mathrm{P}}(10)$	0.806	$\mathrm{m}(x_i^{\max})$	0.638	DFS_y	0.874
10	$r_y^{\mathrm{S}}(10)$	0.803	HjM_y	0.638	$\mathrm{IQR}(t_i^{\max,\min})$	0.868

4. Discussion and conclusion

EEG is a highly complex and dynamic signal. It underlies circadian and ultradian rhythms and changes due to changes in arousal, e.g. sleep changes the characteristics of the EEG, so that EEG is prone to artifacts. The EEG signals depend on the recording setup. Different recording montages will give different spatial filtering and hence different information in the EEG [Nunez, 1995]. The reported findings here are based on EEG with a calculated average reference. It is expected that changing the spatial filtering parameters will give a different result.

A number of earlier studies were optimistic about seizure prediction based on the profile of a measure estimated on segments of the EEG signal, often involving substantial computational complexity. Later studies using larger EEG databases questioned the predictive performance of these same measures and their feasibility in clinical practice. These studies were mostly concentrated on specific measures, such as correlation dimension [Aschenbrenner-Scheibe et al., 2003; Harrison et al., 2005], accumulated energy [Maiwald et al. 2004] and Lyapunov exponents [Lai et al., 2003]. Our study is rather comprehensive with respect to the measures of scalar time series analysis applied to EEG, and all measures are tested on the same EEG data base, that admittedly is of limited size including only 7 preictal records. The study uses only preictal records from scalp EEG without artifact rejection to assimilate clinical practice conditions. The performance of the measures is evaluated in terms of their ability to discriminate different periods of the preictal record, most importantly the late preictal stage (30 min before seizure onset) to earlier preictal stages of the same duration.

On top of the measures earlier used in EEG analysis, we have included other measures that have not been applied to EEG, such as the Spearman and Kendall autocorrelation and the detrended fluctuation analysis, or have been used in EEG analysis very recently, such as the Hurst exponent in [Nurujjaman et al., 2007]. A large proportion of the studied measures are new measures introduced in order to capture different characteristics of the EEG signal and may therefore be useful is seizure prediction, such as the cumulative autocorrelation measures and the measure of deviation of the mutual information from the "Gaussian" mutual information dI_y. Moreover, many new measures are proposed that are either simple statistics of oscillating features or other more advanced measures on the feature time series.

The estimation of all these measures and the comparison of their discriminating power using AUC revealed some interesting points. First, it is confirmed that the level of discrimination of preictal stages depends on physiological conditions. Though we accounted for the seizure type and used records only for generalized tonic clonic seizures, the discrimination performance of the measures varied with the epileptic episode and channel, suggesting that the physiological activity and brain area connectivity varies with the seizure. In opposition to its name, general seizures are not general. Also there is a common notion that the seizures start at different locations even though denoted as primarily generalized seizures. It have been suggested that it is a difference between night and day with start either temporal or frontal in the brain. This is probably oversimplified, but stresses the complexity in the evolution and spreading of epileptic activity in the preictal period.

The results do not show clear discrimination of the late preictal state (last 30min) from earlier periods (from about 3h to 40 min prior to seizure onset) as no measure score consistently high in all seizure episodes, but different measures achieved large AUC values at different seizure episodes. However, overall many measures find significant discrimination of the late preictal state from many earlier preictal stages suggesting that the seizure prediction is at cases possible. This needs further investigation including more episodes as well as records going further backwards in time from the seizure onset into the interictal state.

Besides the limitations of this study in terms of data resources, it is rather clear that simple measures can discriminate preictal stages as well as or even better than measures that are more conceptually advanced and computationally intensive. This was shown already in [Kugiumtzis and Larsson, 2000] for some few measures in small preictal scalp EEG records and was later reported in other works (see [Mormann et al., 2007] for a comprehensive review). Among the highest ranked measures are the median of oscillating features, in particular the inter-spike intervals and local minima and maxima. Even for the seizure episodes they were not among the first best measures, their AUC values were not long behind the first best. These measures are not dependent on the time series length and could be computed on smaller segments, down to a couple of seconds, similarly to the short term Fourier transform. Moreover, they are computationally very efficient and can potentially be used in on-line brain monitoring. Some traditional measures on EEG, i.e. the Hjorth parameters and the energy in specific bands, have scored very high in some cases, but very low in other cases. Correlation measures, such as the Spearman autocorrelation at lag 10 and the decorrelation time, gave also high scores, but they also lack consistent high score in all seizure episodes. It is noted that many measures detecting subtle characteristics in the signal, such as the third order joint moments, the deviation of the mutual information from "Gaussian" mutual information, cross-correlation of oscillating features and models on oscillating features, performed relatively poorly and the same holds for some measures based on nonlinear dynamics theory.

At this stage, many measures are used but on relatively few preictal EEG records, so the results cannot be taken as conclusive. However, it is in the intension of the authors to extent this study and include more EEG preictal records, also under different conditions, i.e. EEG data after removing artefacts, intracranial records, more seizure episodes and other types of seizures.

Acknowledgements

This research project (03ED748) is implemented within the framework of the "Reinforcement Programme of Human Research Manpower" (PENED) and it is co-financed at 90% jointly by E.U.-European Social Fund (75%) and the Greek Ministry of Development -GSRT (25%) and at 10% by Rikshospitalet, Norway.

References

Andrzejak RG, Mormann F, Kreuz T, Rieke C, Kraskov A, Elger CE, Lehnertz K. Testing the null hypothesis of the nonexistence of a preseizure state, *Physical Review E*, 67: 010901, 2003.

Andrzejak RG, Mormann F, Widman G, Kreuz T, Elger CE, Lehnertz K. Improved spatial characterization of the epileptic brain by focusing on nonlinearity, *Epilepsy* 69, 30-44, 2006.

Armitage P, Berry G, Matthews JNS. Statistical Methods in Medical Research (4th edition). Oxford, Blackwell Science, 2002.

Aschenbrenner-Scheibe R, Maiwald T, Winterhalder M, Voss HU, Timmer J, Schulze-Bonhage A. How well can epileptic seizures be predicted? An evaluation of a nonlinear method, *Brain*, 126: 2616-2626, 2003.

Burioka N, Cornélissen G, Halberg F, Kaplan DT, Suyama H, Sako T, Shimizu E. Approximate entropy of human respiratory movement during eye-closed waking and different sleep stages, *Chest*, 123: 80-86, 2003.

Chillemi S, Balocchi R, Di Garbo A, D'Attellis CE, Gigola S, Kochen S, Silva W. Discriminating preictal from interictal states by using coherence measures, *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology* 3: 2319-2322, 2003.

Cover TM, Thomas JA. Elements of Information Theory. New York, Wiley, 1991.

Darbellay G, Vajda I. Estimation of information by an adaptive partitioning of the observation space, *IEEE Transactions on Information Theory* 45 (4), 1999.

Dangel S, Meier PF, Moser HR, Plibersek S, Shen Y. Time series analysis of sleep EEG, *Computer assisted Physics,* 93-95, 1999.

Ehlers CL, Havstad J, Prichard D, Theiler J. Low Doses of Ethanol Reduce Evidence for Nonlinear Structure in Brain Activity, *The Journal of Neuroscience*, 18(18): 7474-7486, 1998.

Gudmundsson S, Runarsson TP, Sigurdsson S, Eiriksdottir G, Johnsen K. Reliability of quantitative EEG features, *Clinical Neurophysiology*, 118: 2162-2171, 2007.

Harrison MAF, Osorio I, Frei MG, Asuri S, Lai Y-C. Correlation dimension and integral do not predict epileptic seizures, *Chaos*, 15: 33106, 2005.

Hegger R, Kantz H, Schreiber T. Practical implementation of nonlinear time series methods: The TISEAN package, *Chaos*, 9: 413-435, 1999.

Hjorth, B. EEG analysis based on time domain properties*. Electroencephalography and Clinical Neurophysiology*, 29: 306-310, 1970.

Iasemidis LD, Sackellares JC. The evolution with time of' the spatial distribution of the largest Lyapunov exponent on the human epileptic cortex, *Measuring Chaos in the Human Brain, (World Scientific, Singapore,)* 49-82. 1991

Iasemidis LD, Shiau D-S, Pardalos PM, Chaovalitwongse W, Narayanan K, Prasad A, Tsakalis K, Carney PR, Sackellares JC. Long-term prospective on-line real-time seizure prediction, *Clinical Neurophysiology,* 116(3): 532-544, 2005.

Kantz H, Schreiber T. Nonlinear Time Series Analysis, Cambridge University Press, Cambridge, 1997.

Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction, *Physical Review A* 45(6), 3403-3411, 1992.

Kugiumtzis D, Larsson PG. Linear and nonlinear analysis of EEG for the prediction of epileptic seizures, “Chaos in Brain?”, Proceedings of the 1999 Workshop, Eds Lehnertz K, Arnhold J, Grassberger P, Elger CE, Singapore, World Scientific, 329-332, 2000.

Kugiumtzis D, Papana A, Tsimpiris A, Vlachos I, Larsson PG. Time series feature evaluation in discriminating preictal EEG states, *Lecture Notes in Computer Science*, 4345: 298-310, 2006.

Lai Y-C, Harrison MAF, Frei MG, Osorio I. Inability of Lyapunov exponents to predict epileptic seizures, *Physical Review Letters*, 91: 068102, 2003.

Lerner DE. Monitoring changing dynamics with correlation integrals: Case study of an epileptic seizure, *Physica D: Nonlinear Phenomena* 97 (4): 563-576, 1996.

Maiwald T, Winterhalder M, Aschenbrenner-Scheibe R, Voss HU, Schulze-Bonhage A, Timmer J. Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic, *Physica D*, 194: 357-368, 2004.

Martinerie J, Adam C, Le Van Quyen M, Baulac M, Clemenceau S, Renault B, Varela FJ. Epileptic seizures can be anticipated by non-linear analysis, *Nature Medicine* 4 (10), 1998.

McSharry PE, Smith LA, Tarassenko L. Comparison of predictability of epileptic seizures by a linear and a nonlinear methods**,** *IEEE Trans on BME 2003,* 50(5)**:**628-633, 2003.

Mormann F, Kreuz T, Rieke C, Andrzejak R, Kraskov A, David P, Elger C, Lehnertz K. On the predictability of epileptic seizures, *Clinical Neurophysiology*, 116: 569-587, 2005

Mormann F, Elger CE, Lehnertz K. Seizure anticipation: from algorithms to clinical practice, *Current Opinion in Neurology*, 19(2):187-193, 2006.

Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: The long and winding road, *Brain* 130:314-333, 2007.

Nicolaou N, Nasuto SJ. Mutual Information for EEG analysis, IEEE EMBSS UKRI Postgraduate Conference on Biomedical Engineering and Medical Physics (PGBIOMED'05), 2005.

Nielsen H.A, Madsen H. A generalization of some classical time series tools, *Computational Statistics & Data Analysis* 37, 13-31, 2001.

Nurujjaman Md, Naryanan R, Iyengar ANS. Comparative study of nonlinear properties of EEG signals of a normal person and an epileptic patient, *arXiv:0709.3683*, 2007.

Papana A, Kugiumtzis D. Evaluation of mutual information estimators on nonlinear dynamic systems, submitted to *Nonlinear Phenomena in Complex Systems*, 2007.

Peng CK, Havlin S, Stanley HE, Goldberger AL. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time serie, *Chaos*, 5:82-87, 1995.

Sackellares JC, Shiau DS, Iasemidis LD, Dance LK, Pardalos PM and Chaovalitwongse W. Multi-Dimensional Multi-Parameter Time Series Processing for Seizure Warning and Prediction. Patent number US 7,263,467, Date of Patent: Aug. 28, 2007

Ziv J, Lempel A. Compression of individual sequences via variable-rate coding, *IEEE Transactions for Information Technology,* 24(5): 530-536, 1978.