



ELSEVIER

Physica A 300 (2001) 505–520

PHYSICA A

www.elsevier.com/locate/physa

Nonlinear structures in electroencephalogram signals

L. Diambra^{a,*}, C.P. Malta^b, A. Capurro^c, J. Fernández^d

^a*Instituto de Ciências Biomédicas, Universidade de São Paulo, 05508-900 São Paulo, Brazil*

^b*Instituto de Física, Universidade de São Paulo, C.P. 66318, 05315-970 São Paulo, Brazil*

^c*Instituto de Investigaciones Biológicas Clemente Estable, Av. Italia 3318, Montevideo, Uruguay*

^d*Center for Neurobiology and Behavior, Columbia University, 722 W. 168th Street, New York, NY 10032, USA*

Received 20 December 2000; received in revised form 22 May 2001

Abstract

We apply a nonlinear prediction algorithm to investigate the presence of nonlinear structure in electroencephalogram (EEG) recordings. The EEG signal could be modeled as a realization of a nonlinear model plus a residual noise (uncorrelated Gaussian noise). Using linear and nonlinear models we analyze the statistical nature of these residual noises in the case of epileptic patients and normal subjects. We found that the residual noise presents Gaussian distribution for epileptic patients if a nonlinear model is used whereas in the case of normal subjects the residual noise will exhibit a Gaussian distribution only if a linear model (autoregressive) is used. These results provide another evidence of the nonlinear character of the epileptic seizure recordings, while the normal EEG seems to be better described as linearly correlated noise. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Electroencephalogram; Epilepsy; Nonlinear models

1. Introduction

It is a nontrivial problem to decide whether a given time series is adequately described by a linearly filtered noise or contains nonlinearities (deterministic chaos) [1–6]. The situation becomes even more complicated if the time series is non-stationary, and results from a combination of nonlinearities and random perturbations as it seems to be the case of electroencephalogram (EEG) signals (for review see

* Corresponding author.

E-mail address: diambra@fisio.icb.usp.br (L. Diambra).

Refs. [7,8]). In this paper, we shall test three different models for identifying nonlinear structures in the dynamics underlying the EEG signals recording.

During recent years several neurophysiological studies have shown that EEG signals have high degree of complexity. This complexity may be due to random processes or to the chaotic behavior of the underlying nonlinear dynamics [9]. In fact, many authors have reported evidence that epileptiform EEG signals may exhibit low dimensional deterministic chaos [10–14]. This means that the dynamical evolution of the system can be described with only few variables, and one can construct a model able to predict the short time future behavior of the system in terms of its previous states [15]. On the other hand, normal EEG signals have higher complexity, and they could not be distinguished from linearly filtered noise [2,3], with the possible exception of occipital recordings [5]. Up to now, the bulk of these studies are based on measurements of some metric parameters like the correlation dimension [16] or Lyapunov exponent [17]. These methods may easily produce spurious interpretation when care is not taken. For this reason, other methods that are not based on invariant measures are welcome for the characterization of different types of dynamics in EEG signals.

The use of other quantitative measures for comparing aspects of the EEG in different conditions has become an increasingly effective EEG research tool [18–20]. Recently, studies based on non-linear analysis have been able to anticipate epileptic seizures from intracranially recorded EEG [21,22]. In the present effort, the main goal is the identification of nonlinear structures in the EEG signals. We examine the possibility of EEG signals being modeled basically by a nonlinear deterministic model related to the activity of a population of neurons highly correlated in their discharging time, plus some level of residual noise resulting from discharges of a more desynchronized population of neurons. Consequently, we test as null hypothesis that residual noise (RN) from linear and nonlinear models are realizations of a Gaussian random process. The RN is defined as the prediction error, i.e., the difference between the EEG signal and the prediction of a given model.

In order to build nonlinear prediction models for the dynamics underlying the highly correlated activity, we use a parametrized nonlinear function, and the parameters are determined, within the framework of the approach presented in Ref. [18], using a short portion of EEG. This approach allows the determination of the parameters using around 450 points (in an epileptic regimen). Thus, we can determine the embedding dimension of short portions of EEG by estimating the ability of prediction of the model constructed with it. In this way, problems associated with the non-stationarity of the signal are avoided. We found that the embedding dimension estimated using this method is in agreement with estimates provided by the false nearest-neighbor (FNN) method in the case of low-dimensional dynamics.

We organize our presentation as follows: in Section 2, we describe the procedure used to record the EEG signals, and review some ideas concerning the reconstruction of the system's state and parameters estimation procedure based on the maximum entropy principle (MEP). In Section 3, we present our results regarding the EEG signals from

normal subjects during wakefulness and epileptic patients during inter-ictal periods. Finally, some conclusions are drawn in Section 4.

2. Materials and methods

2.1. Clinical data

As control group we have taken the EEG recordings from three normal individuals, males, between 26 and 47-years old, free of a past history or current symptoms of psychopathology. The recordings have been obtained under wakefulness, with the individuals physically and mentally at rest, and closed eyes. For the epileptic group, we considered EEG recordings from four males and two females, between 5 and 39-years old. Three of the patients presented clinical history of generalized seizure, and three of partial seizure that turned into generalized. The recordings were performed during inter-ictal periods. The epileptic patients were examined under one of the following conditions: in the control condition of the normal patients, spontaneous or induced sleep, and hyperventilation. The last two conditions are usually employed in clinical recordings in order to induce the appearance of the epileptic inter-ictal activity. These EEG recordings were examined by a clinical specialist in order to select segments containing evident paroxysmal discharges and free of artifacts.

The recordings have been obtained using a standard clinical device (International 10/20 Systems) connected to the scalp, a reference electrode being placed at the patient's nose. We used the time series obtained from frontal derivations to apply our methods. The data from epileptic patients were amplified and filtered using a low-frequency cutoff of 0.1 Hz, and high-frequency cutoff of 50 Hz. The data were stored on magnetic tape and then digitized off-line at 256 Hz with 8-bit digitizer.

2.2. Nonlinear modeling

The complexity of an attractor can be characterized by a set of invariants, the *generalized dimension spectrum*. Generally speaking, for a reliable estimation of these invariants, large quantities of data are necessary to achieve accurate approximation for the density of points in different regions of the attractor. One needs a long stationary time series. This condition may not be satisfied when we deal with EEG signals [24,25], since only few seconds of stationary-like recordings are usually available.

Here, we assume that EEG signals can be basically described as deterministic chaos with some level of additive random noise [26]. We need to reconstruct the space state associated with the deterministic part of the signal. The EEG signals are a stroboscopic sequence of N measurements $\{v(t_0), v(t_0 + \tau_s), \dots, v(t_0 + N\tau_s)\}$ made at intervals τ_s . The state space is reconstructed using the time delay embedding [27,28], which uses a collection of coordinates with time lag to create a vector in d dimensions,

$$\mathbf{v}(t) = (v(t), v(t - \Delta), \dots, v(t - (d - 1)\Delta)), \quad (1)$$

where $\Delta = n\tau_s$, (with $n = 1, 2, \dots, N$), is the time lag or delay. Takens has shown [27] that for flows evolving to compact attracting manifolds of dimension d_a , if $d > 2d_a$, we can write

$$v(t + T) = F(\mathbf{v}(t)), \quad (2)$$

where $T > 0$ is the forecasting time. This means that the dynamical evolution of the system could be described with few variables and one can construct a model $F^*(\mathbf{v}(t))$ that is able to predict approximately the short future behavior of the system in terms of its previous states. The Takens' theorem provides no information regarding either the choice of Δ , or the form of F .

Now, we introduce a parametrized mapping function $F^*(\mathbf{v}(t), \mathbf{a})$, where \mathbf{a} is the set of parameters of the model [23]. These parameters are determined within the framework of the approach presented in [18,23]. We consider here three types of representation of the mapping function. In the first case we consider a linear model (labeled model A) $F^*(\mathbf{v}(t), \mathbf{a}) = F_A(\mathbf{v}(t), \mathbf{a}_A)$, given by

$$F_A(\mathbf{v}(t), \mathbf{a}_A) = a_0 + \sum_{j=1}^d a_j v_j, \quad (3)$$

where v_j are the components of the d -vector (1). The working hypothesis is given by the set of parameters $\mathbf{a}_A = \{a_0, a_1, \dots, a_d\}$.

A more general predictor form is provided by a kind of radial basis expansion (labeled model B) $F^*(\mathbf{v}(t), \mathbf{a}) = F_B(\mathbf{v}(t), \mathbf{a}_B)$, where $F_B(\mathbf{v}(t), \mathbf{a}_B)$ is given by the expansion

$$F_B(\mathbf{v}(t), \mathbf{a}_B) = \sum_{j=1}^d \sum_{i=1}^g a_{ij} \exp \left[-\frac{g}{2} (v_j - x_i)^2 \right]. \quad (4)$$

We use g Gaussian functions centered at x_i . The number of Gaussian functions was determined using the criterion of best fitting, i.e., the number of functions is gradually increased until no improvement in the predictive performance is observed. This was achieved for $g = 6$. Then, we normalize the signal to the unity by dividing all data points by their maximum absolute value. This normalization is necessary in order to have the same scale for all patients. The x_i are equidistant points¹ in the interval $[-1.5, 1.5]$. This interval is chosen in order to guarantee a good fitting in the interval of interest, i.e., $[-1, 1]$. Of course, the a_{ij} constitute our working hypothesis, $\mathbf{a}_B = \{a_{1,1}, a_{2,1}, \dots, a_{g,d-1}, a_{g,d}\}$. Thus, the number of parameters of the model, N_B , is determined by the number g of Gaussian functions in (4), and by the embedding dimension d , $N_B = d \times g$. Expansion (4) provides a smooth interpolation of scattered data in arbitrary dimension [15], and has proven to be useful in practice [18].

We will use also a quadratic expansion $F^*(\mathbf{v}(t), \mathbf{a}) = F_C(\mathbf{v}(t), \mathbf{a}_C)$ where the quadratic predictor $F_C(\mathbf{v}(t), \mathbf{a}_C)$ is given by

$$F_C(\mathbf{v}(t), \mathbf{a}_C) = a_0 + \sum_{j=1}^d a_j v_j + \sum_{j=1}^d \sum_{i=1}^j a_{i,j} v_i v_j. \quad (5)$$

¹ We use in this work $g = 6$, so that $x_i = 3i/6 - 1.5$, with $i = 0, 1, \dots, 5$.

In this case, the working hypothesis is given by the set of parameters $\mathbf{a}_C = \{a_0, a_1, a_2, \dots, a_d, a_{1,1}, \dots, a_{d,d}, a_{1,2}, a_{1,3}, \dots, a_{d-1,d}\}$. If one takes the first two terms in (5), we have the linear model (3) or the autoregressive model of order d . Notice that the parameters of the models A, B, C have been arranged so as to be represented by the row vector \mathbf{a}_A , \mathbf{a}_B , \mathbf{a}_C , respectively. From now on we shall use the row vector \mathbf{a} to denote a generic set of parameters ($\mathbf{a} = \mathbf{a}_A$ in the case of model A, $\mathbf{a} = \mathbf{a}_B$ in the case of model B, $\mathbf{a} = \mathbf{a}_C$ in the case of model C).

Now we introduce the MEP [18], i.e., the minimum number of assumptions compatible with the available data, in order to determine the parameters \mathbf{a} using the information contained in a segment of $K = M + d\Delta$ points of the EEG time series. These points are required to construct M sets of the type

$$\{\mathbf{v}(t_j), v(t_j + T)\}, \quad j = 1, \dots, M, \tag{6}$$

where j are labeling the M sets. In order to infer the coefficients consistent with these M sets (6), our central idea is to assume that *each set of parameters \mathbf{a} is realized with probability $P(\mathbf{a})$* , and reinterpret the M sets (6) according to expression (2), where the parameters are set equal to the mean value of the distribution $P(\mathbf{a})$,

$$\begin{aligned} v(t_j + T) &= F^*(\mathbf{v}(t_j), \mathbf{a})|_{\mathbf{a}=\langle \mathbf{a} \rangle} \\ &= \mathbf{V}_j \cdot \langle \mathbf{a} \rangle^t, \end{aligned} \tag{7}$$

where \mathbf{V}_j are row vectors constructed by evaluating the function component of series (3)–(5) in the d -vector $\mathbf{v}(t_j)$,

$$\mathbf{V}_j = \{1, v(t_j), v(t_j - \Delta), \dots, v(t_j - (d - 1)\Delta)\} \tag{Model A ,}$$

$$\mathbf{V}_j = \{e^{-(g/2)(v(t_j)-x_1)^2}, e^{-(g/2)(v(t_j)-x_2)^2}, \dots, e^{-(g/2)(v(t_j)-(d-1)\Delta-x_g)^2}\} \tag{Model B ,}$$

$$\begin{aligned} \mathbf{V}_j &= \{1, v(t_j), v(t_j - \Delta), \dots, v^2(t_j), v^2(t_j - \Delta), \dots, v(t_j - (d - 2)\Delta) \\ &\quad \times v(t_j - (d - 1)\Delta)\} \end{aligned} \tag{Model C}$$

and $\langle \mathbf{a} \rangle^t$ is the transpose of the row vector having as components the expectation values

$$\langle a_i \rangle = \int P(\mathbf{a}) a_i \, d\mathbf{a} . \tag{8}$$

The two essential ingredients of the MEP technique are the expectation values $\langle a_i \rangle$ (8), and the *relative entropy* S associated with the probability distribution $P(\mathbf{a})$ [29,30] defined by

$$S = - \int P(\mathbf{a}) \ln \left(\frac{P(\mathbf{a})}{P_0(\mathbf{a})} \right) \, d\mathbf{a} , \tag{9}$$

where $P_0(\mathbf{a})$ is an a priori distribution. Here we select a Gaussian distribution as P_0 . As customary, one is then led to maximizing entropy (9) subject to constraints (7) and the normalization condition, obtaining

$$S' = - \int \left\{ P(\mathbf{a}) \ln \left(\frac{P(\mathbf{a})}{P_0(\mathbf{a})} \right) + \lambda_0 P(\mathbf{a}) + \mathbf{a} \cdot \mathbf{W}^t \vec{\lambda} P(\mathbf{a}) \right\} \, d\mathbf{a} , \tag{10}$$

where λ_0 and the column vector $\vec{\lambda}$ are Lagrange multipliers associated, respectively, with the normalization condition and constraints (7), and \mathbf{W} is a matrix with M rows \mathbf{V}_j . Variation of S' with respect to $P(\mathbf{a})$ immediately gives

$$P(\mathbf{a}) = \exp(-(1 + \lambda_0)) \exp(-\mathbf{a} \cdot \mathbf{\Gamma}) P_0(\mathbf{a}), \quad (11)$$

where $\mathbf{\Gamma} = \mathbf{W}^t \vec{\lambda}$ (\mathbf{W}^t is the transpose of \mathbf{W}).

A choice has now to be made concerning the a priori probability distribution P_0 . Here we select a Gaussian P_0 , i.e., choose it to be proportional to $\exp(-\mathbf{a} \cdot \mathbf{a} / 2\sigma)$, with a free parameter σ . When we replace this choice of the a priori distribution in Eq. (11), we obtain a Gaussian form for the probability distribution $P(\mathbf{a})$, centered at $\langle \mathbf{a} \rangle = -\sigma \mathbf{\Gamma}$, with dispersion σ . Both the definition of $\mathbf{\Gamma}$, and constraints (7) allow for the elimination of the Lagrange multipliers $\vec{\lambda}$. One can thus express the $\langle a_i \rangle$, solely in terms of the data set:

$$\langle \mathbf{a} \rangle^t = I_{ps}[\mathbf{W}] \mathbf{v}_T, \quad (12)$$

where $I_{ps}[\mathbf{W}] = (\mathbf{W})^t (\mathbf{W}(\mathbf{W})^t)^{-1}$ is the Moore–Penrose pseudoinverse [31], and \mathbf{v}_T is a column vector constructed with $v(t_j + T)$, $j = 1, \dots, M$. Now, we choose the most probable set of parameters (the mean value of the distribution) compatible with constraints (7) as our working hypothesis \mathbf{a} . Using the above procedure, one can capture the dynamics underlying a short portion of the EEG, i.e., underlying the M sets (6).

2.3. Procedure outline

We aim at applying this methodology to EEG signals regarding two situations: normal EEG during wakefulness, and during epileptic inter-ictal activity in epileptic patients. The first step is to select one quasi-stationary segment of the EEG from each patient. Then we will determine the adequate set of embedding parameters (time delay Δ and the embedding dimension d) for each model and each patient using the first $K = M + d\Delta$ points of the selected segment. In order to obtain the RN, we test the three models (linear model A, and the two nonlinear models B, C) over each segment.

We choose as time delay embedding Δ the first minimum (τ_0) of the mutual information [32], for each segment to be modeled. In Fig. 1 (thick line) we show the mutual information of the EEG signal from the epileptic patient #1176. In order to select the adequate embedding dimension, we used two criteria. First, we analyzed the predictive performance of the models in a segment of EEG with K points as a function of the embedding dimension, and then we chose the optimal predictor (labeled as criterion I). The forecasting time T is set equal to the corresponding delay embedding Δ used in the embedding procedure. We say that a predictor is optimal, for a given time series, if the mean square error of the RN is minimal [33]. The mean square error δ is computed over the M sets (6):

$$\delta = M^{-1} (\mathbf{v}_T - \mathbf{v}_T^*) \cdot (\mathbf{v}_T - \mathbf{v}_T^*), \quad (13)$$

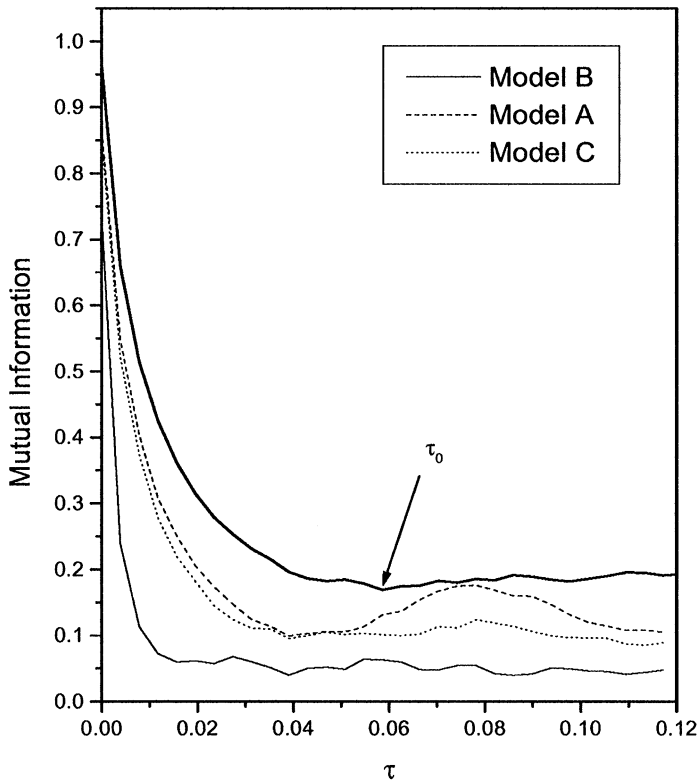


Fig. 1. Mutual information of the EEG signal from the epileptic patient #1176 (thick line). The first minimum τ_0 of the mutual information function occurs at 0.058 s (15 samples). Mutual information of the RN for the patient #1176 using criterion II: linear predictor (model A) in dashed line; Gaussian function expansion (model B) in solid line; and quadratic predictor (model C) in dotted line.

where \mathbf{v}_T corresponds to the EEG signal and \mathbf{v}_T^* to the model prediction. In Fig. 2, we show the mean square error δ versus the embedding dimension d using the nonlinear model B constructed with three different values of the time delay Δ ($\Delta = 14$ samples in solid line, $\Delta = 15$ samples in dashed line, and $\Delta = 16$ samples in dotted line). We can see that there is a minimum when the embedding dimension reaches $d = 3$. This value depends on the signal nature, and it is lower in EEG segments with epileptic activity than in EEG segments from normal subjects. In order to perform our calculation, we used 400 sets ($M = 400$) for modeling epileptic activity, and 600 sets ($M = 600$) in the case of normal subjects (remember that the number of points K is given by $M + d\Delta$). We choose the dimension at which the minimum occurs as the embedding dimension for the modeling. In order to check the embedding dimension suggested by the criterion I, we also estimate the embedding dimension by the method of false nearest neighbor [34] (labeled as criterion II). In low dimension cases (epileptic subjects) the two criteria have given a similar estimate for the embedding dimension (see Fig. 2), but there is

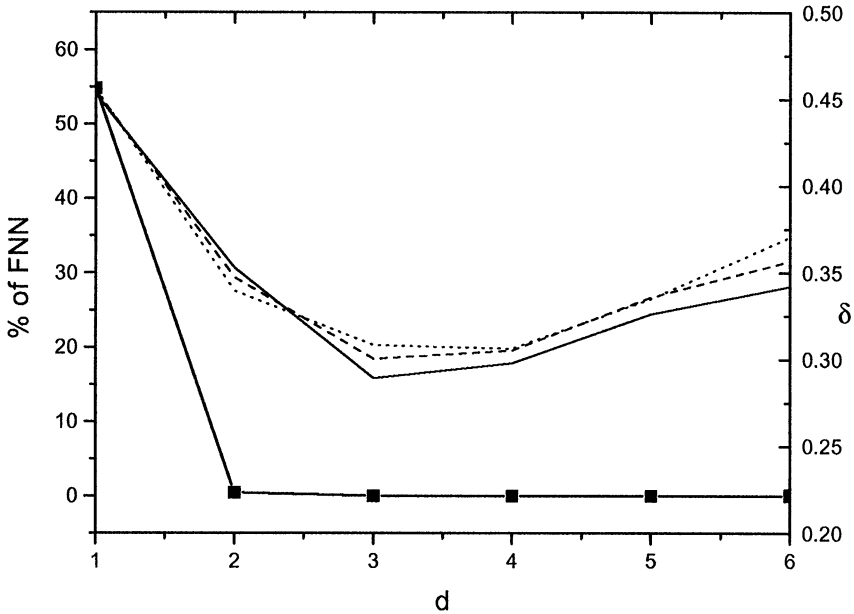


Fig. 2. Mean deviation δ (right axis) as a function of the embedding dimension d for the Gaussian function expansion (4) (model B), from EEG segment with epileptic activity of the patient #1176. We used three time delay in the embedding: 14 samples (solid line), 15 samples (dashed line), and 16 samples (dotted line). The models were constructed using 400 points. For the sake of comparison, we also show (left axis) the percentage of false nearest neighbors (thick line) as function of d .

a great divergence in the case of normal subjects (see Fig. 3). In particular, we found that the embedding dimension suggested by FNN is smaller than that suggested by criterion I. This underestimation could be due to the small quantity of points used for estimating the percentage of false nearest neighbor (around 1500 data points). The threshold parameter used for estimating false nearest neighbor was the standard deviation of the data divided by 7.

2.4. Null hypothesis

Once the embedding parameters are set, we test each model against a null hypothesis. In this paper the null hypothesis consists of two independent parts: (i) the RN is normally distributed, and (ii) the RN is independent. In order to test the normality of the RN distribution, we constructed a histogram for the RN of the selected segment. The number of bins used in the construction of the histograms lies in the range between 30 and 45, the larger the number of points in the RN, the larger the number of bins used (we have used 1500 points in the case of normal subjects, and in the case of epileptic patients we used between 650 and 850 points, depending on the length of the quasi-stationary segment considered). Then the normality of the distribution of the RN is formally tested via the well-known χ^2 method. In order to test if the RN is

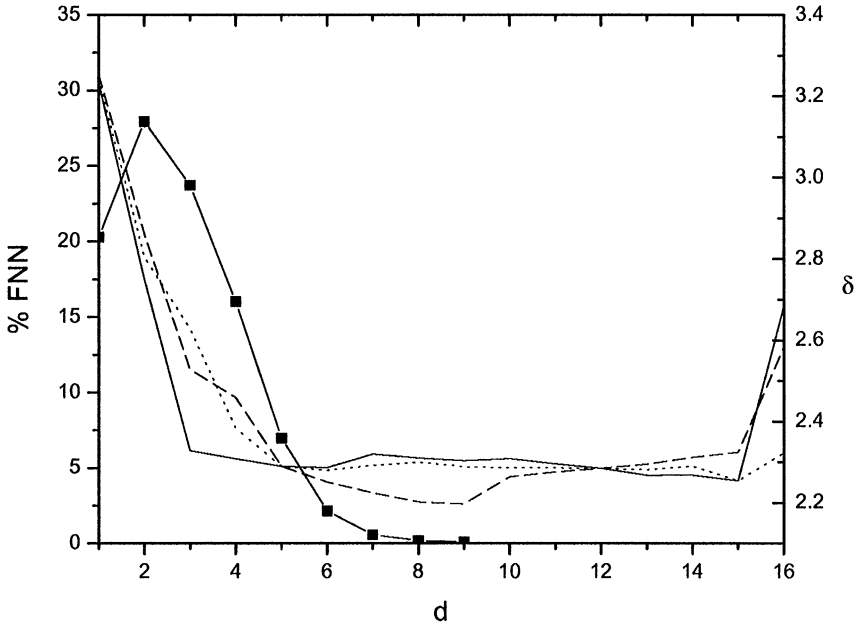


Fig. 3. Mean deviation δ (right axis) as a function of the embedding dimension d for the Gaussian function expansion (4) (model B), from EEG segment from normal subject #1003. We used three time delay in the embedding: 14 samples (solid line), 15 samples (dashed line), and 16 samples (dotted line). The models were constructed using 1000 points. For the sake of comparison, we also show (left axis) the percentage of false nearest neighbors (thick line) as function of d .

statistically independent we estimated the mutual information function. Given a time series $\{s_1, s_2, \dots, s_J\}$ the mutual information is defined as

$$I_{\text{Mutual}}(\tau) = \sum_{i=1}^J P(s_i, s_{i+\tau}) \log \frac{P(s_i, s_{i+\tau})}{P(s_i)P(s_{i+\tau})}. \tag{14}$$

Thus, $I_{\text{Mutual}}(\tau)$ gives the degree of statistical autocorrelation, and $I_{\text{Mutual}}(\tau) = 0$ if s_i and $s_{i+\tau}$ are statistically independent.

3. Results

We tested the performance of the three models in each portion of EEG: the linear model A, and the two nonlinear models B, C. In Table 1, we show the time delay embedding Δ , and the embedding dimension d , used in the modeling procedure of one portion of each patient EEG.

We tested the null hypothesis in the RN from six epileptic patients. In Table 2, we show the p -values for rejecting the null hypothesis for the RN of the three models being considered, using the embedding dimension suggested by criterion I. In the case of the linear model A, the null hypothesis is rejected without ambiguity for all EEG

Table 1

Embedding dimension and time delay embedding Δ used for modeling each EEG segment of the patients in the first column. The last three rows correspond to normal subjects. We have not observed any minimum or plateau in the case of EEG of normal subjects modeled with the model C, so we did not perform the calculations for this case

Patient	Embedding dimension				Δ in samples
	Model A	Model B	Model C	FNN	
1072 F3	9	4	4	3	21
1176 F4	8	3	5	2	15
0176 F4	8	3	6	3	15
0748 F4	4	3	4	3	25
0857 F4	8	5	6	3	14
0994 F4	5	3	4	4	30
1001 F4	25	15	—	6	10
1002 F4	39	13	—	6	12
1003 F4	24	9	—	7	15

Table 2

Criterion I: p -values for rejecting the Gaussian hypothesis for the RN from the three models used. We give the first minimum τ_0 of the mutual information function of the RN in each case. The last three rows correspond to normal subjects

Patient	Model A		Model B		Model C	
	p -value	τ_0	p -value	τ_0	p -value	τ_0
1072 F3	1.0000	0.027	0.0797	0.008	1.0000	0.020
1176 F4	0.9240	0.043	0.2404	0.011	0.9548	0.043
0176 F4	1.0000	0.043	0.1087	0.011	1.0000	0.055
0748 F4	1.0000	0.059	0.0129	0.016	1.0000	0.059
0857 F4	1.0000	0.047	0.1350	0.016	0.8825	0.039
0994 F4	0.9961	0.068	0.0067	0.019	0.8038	0.058
1001 F4	0.0842	0.029	1.0000	0.020	—	—
1002 F4	0.0747	0.039	0.9997	0.039	—	—
1003 F4	0.0161	0.029	0.9247	0.029	—	—

segments from epileptic patients. On the other hand, when we use the nonlinear model B, we find that the null hypothesis is accepted with more than 90% level of significance in three patients, and in two other patients with more than 85%. One patient (#1176) presents a level of significance slightly lower. We can see in Fig. 4 three histograms of amplitudes of the RN from the EEG of this patient using criterion I (embedding dimension = 3). When we tested the null hypothesis with the nonlinear model C, it could be rejected without ambiguity in all cases.

We also tested the null hypothesis in the RN from models constructed using the embedding dimension suggested by criterion II. In Table 3, we show the p -values for rejecting the null hypothesis for the RN of the three models. There is agreement with the results listed in Table 2 in the case of models A and C. In the case of model B, we have totally different results for the patient #1072. For patients #0857

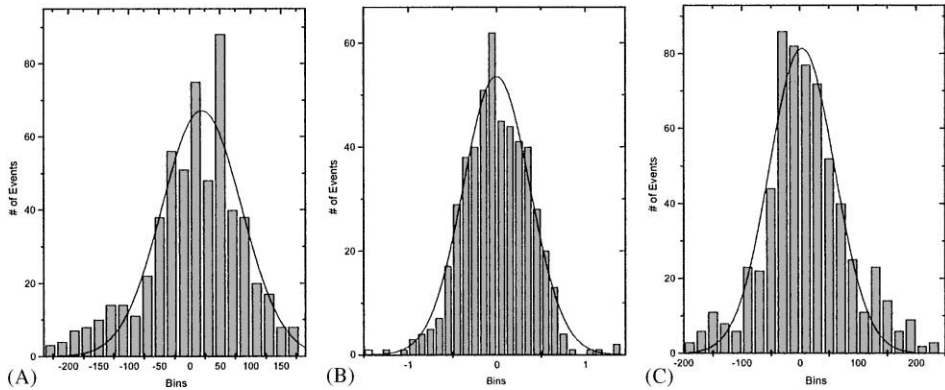


Fig. 4. Histogram and Gaussian fitting of the three residual noise models, using criterion I, for the patient #1176: linear predictor (A); Gaussian function expansion (B); and quadratic predictor (C).

Table 3

Criterion II: p -values for rejecting the Gaussian hypothesis for the RN from the three models used. We give the first minimum τ_0 of the mutual information function of the RN in each case. The last three rows correspond to normal subjects

Patient	Model A		Model B		Model C	
	p -value	τ_0	p -value	τ_0	p -value	τ_0
1072 F3	1.0000	0.082	1.0000	0.023	1.0000	0.046
1176 F4	1.0000	0.043	0.0108	0.011	1.0000	0.043
0176 F4	1.0000	0.047	0.1087	0.011	1.0000	0.043
0748 F4	1.0000	0.047	0.0129	0.016	1.0000	0.055
0857 F4	0.7045	0.051	0.3864	0.016	0.8470	0.039
0994 F4	0.6696	0.146	0.3011	0.029	1.0000	0.146
1001 F4	1.0000	0.029	1.0000	0.029	1.0000	0.029
1002 F4	1.0000	0.039	1.0000	0.020	1.0000	0.039
1003 F4	1.0000	0.039	1.0000	0.020	1.0000	0.039

and #994 the null hypothesis could be accepted with ambiguity. There is agreement in the embedding dimension suggested by both criteria in the case of patients #0176 and #0748. Notice that in the case of the patient #1176 (shown in Fig. 4), the null hypothesis is accepted with more than 95% level of significance. In Fig. 5, we show the histograms of amplitudes for the RN from the EEG of the patient #1176. Only Fig. 5B shows a histogram that is well fitted by a Gaussian function.

From Tables 2 and 3 we see that the first minimum τ_0 of the mutual information function of the RN is smaller than the Δ of the corresponding EEG signal listed in Table 1. In Fig. 1, we can see that the mutual information function of the RN from the EEG of the patient #1176 in the case of model B decreases more rapidly than in the case of models A and C. These results are compatible in general with the idea that the epileptic EEG signals can be generated by a nonlinear system of low dimension plus random fluctuations in the form of Gaussian white noise.

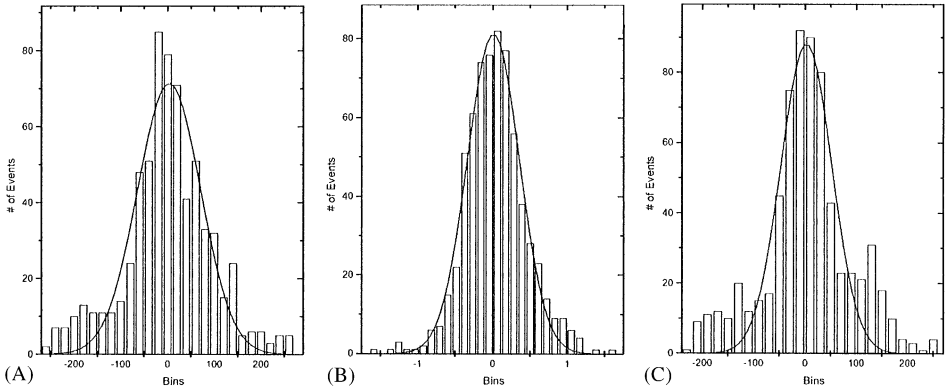


Fig. 5. Histogram and Gaussian fitting of the three residual noise models, using criterion II, for the patient #1176: linear predictor (A); Gaussian function expansion (B); and quadratic predictor (C).

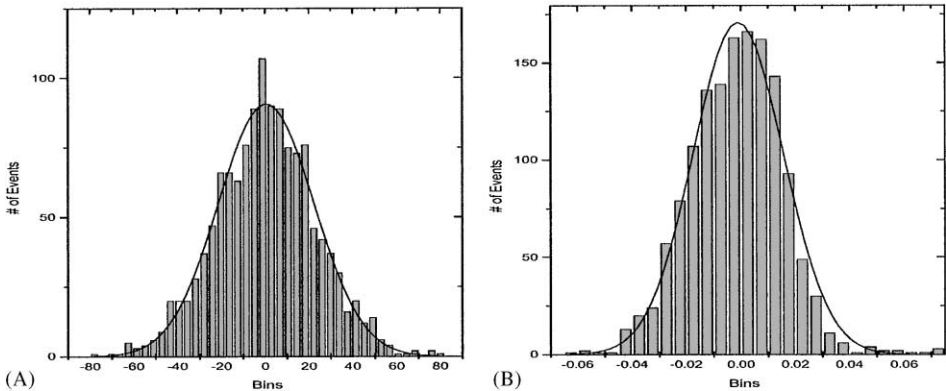


Fig. 6. Histogram and Gaussian fitting of the two residual noise models for the normal subject #1002: linear predictor (A) and Gaussian function expansion (B).

When we deal with EEG signals from normal subjects in wakefulness, the situation is rather different. Let us consider first the RN from model B constructed using the embedding dimension suggested by criterion I. The histogram of the RN from the nonlinear models B (4) does not present a normal distribution. Table 2 shows that the null hypothesis is rejected without ambiguity in the case of the normal subject #1001, with a significance level smaller than 10^{-3} for the normal subject #1002, and with a significance level smaller than 0.1 in the case of the normal subject #1003. However, we tested the RN from the linear model for normality, and we found that the null hypothesis can be accepted in the three cases analyzed with more than 90% level of significance. In Fig. 6, we show the typical histograms of the RN from a normal patient in wakefulness for models A and B. However, as shown in Fig. 7, the values of τ_0 of the RN are not significantly different from the Δ of the corresponding EEG signal. In the case of RN from models constructed using the embedding dimension

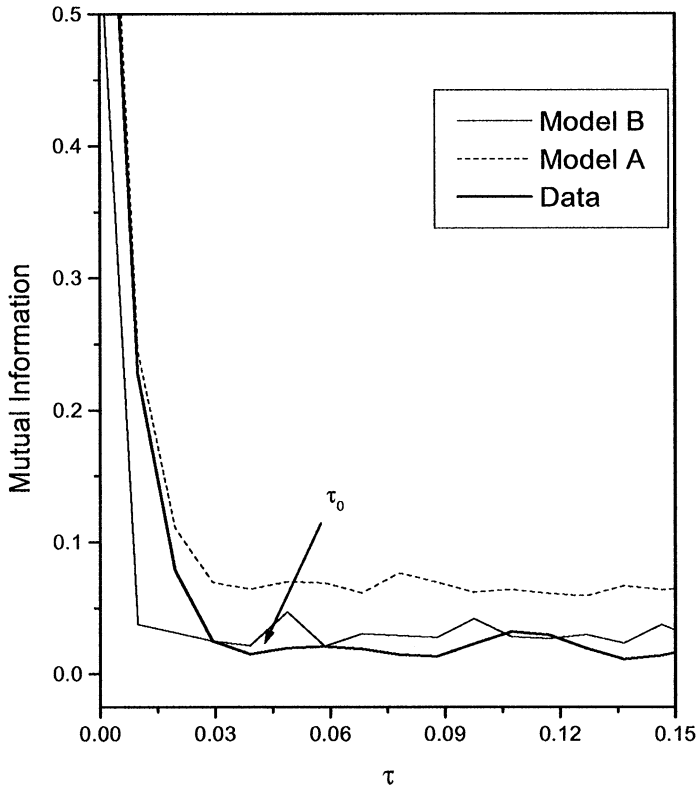


Fig. 7. Mutual information of the EEG signal from the control subject #1002 (thick solid line). The first minimum τ_0 of the mutual information function occurs at 0.049 s. Mutual information of the two RN from the normal subject #1002: linear predictor (model A) in dashed line, and Gaussian function expansion (model B) in solid line.

suggested by criterion II, the null hypothesis is rejected without ambiguity for all models here considered (see Table 3). This could be explained by an underestimation of the embedding dimension due to the limited amount of data used to compute the false nearest neighbors percentage in the high-dimension cases.

4. Discussion and conclusion

In this paper, we consider the EEG to be a realization of two components: a chaotic attractor and Gaussian noise. We consider here only EEG recordings obtained using scalp connections (i.e., of the type that are currently used for standard medical diagnosis). The limited number of cases that we have studied does not allow us to draw definitive conclusions, nevertheless, our results indicate that it is possible to approximate the dynamics of the epileptic EEG signal by a nonlinear oscillator with Gaussian noise. This provides a new evidence suggesting the presence of nonlinear structure

in these signals [3,4,9,10]. This conclusion, based on the Gaussian distribution and independence of the RN sequence, holds only for signals from epileptic subjects. In the case of the frontal EEG recordings of normal subjects in wakefulness we found that the RN from high-dimensional linear model has a Gaussian distribution. However, in this case our method does not allow to conclude that the RN is independent. We think that this problem could be better assessed using a more sophisticated analysis like [35,36].

The embedding dimension of the models are lower when the EEG signal contains epileptic activity. This is in accordance with previous studies that concluded that the complexity of the epileptic EEG is lower than the complexity of the normal one [10–12,20]. However, we would like to stress that this fact does not constitute a rigorous mathematical argument *pro* low-dimensional chaotic dynamics in the epileptic EEG.

As a final comment, we want to outline the hypothesis that each component of the model may be produced by the activity of a different dynamical system, which in turn may correspond to a different population of neurons. In the case of epileptic activity, the chaotic attractor may be produced by the activity of a population of neurons that fire with a high degree of synchronization. It is known that epileptic spikes are due to the abnormal hypersynchronization of big populations of neurons [25,37]. The component of white noise may be due to the activity of a second large population of neurons that have not been recruited in the epileptic spikes thus having a less synchronized discharging regime.

In conclusion, we propose a model to fit epileptic EEG signals that consists of a deterministic system plus white Gaussian noise. The limitations of our conclusion to epileptic EEG signals may be related to the high degree of complexity of the EEG signals from normal patients recorded during wakefulness. In these cases, although we have not been able to determine uncorrelation in the RN, it seems that our results support the hypothesis that this kind of signal can be described in terms of high-dimension linear models.

Acknowledgements

L.D. acknowledges financial support from FAPESP (Brazil), C.P.M. acknowledges the partial financial support from CNPq (Brazil), A.C. acknowledges the financial support of PEDECIBA (Uruguay), and J.F. acknowledges the financial support of UNLP (Argentina).

References

- [1] M. Paulus, Testing for nonlinearity in the EEG, in: B.H. Hansen, M.E. Brandt (Eds.), *Nonlinear Dynamical Analysis of the EEG*, World Scientific, Singapore, 1993, pp. 100–115.

- [2] J. Theiler, P.E. Rapp, Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram, *Electroenceph. Clin. Neurophysiol.* 98 (1996) 213–222.
- [3] J.P.N. Pijn, J. Van Neerven, A. Noest, F. Lopes da Silva, Chaos or noise in the EEG signals; dependence on state and brain site, *Electroenceph. Clin. Neurophysiol.* 79 (1991) 371–381.
- [4] M.J. van der Heyden, C. Diks, J.P. Pijn, D.N. Velis, Time reversibility of intracranial human EEG recordings in mesial temporal lobe epilepsy, *Phys. Lett. A* 216 (1996) 283–288.
- [5] F.H. Lopes da Silva, J.P. Pijn, D.N. Velis, P.C.G. Nijssen, Alpha rhythms: noise, dynamics and models, *Int. J. Psychophysiol.* 26 (1997) 237–249.
- [6] M. Barahona, C.S. Poon, Detection of nonlinear dynamics in short, noisy time series, *Nature* 381 (1996) 215–217.
- [7] E. Niedermeyer, F. Lopes da Silva, *Electroencefalography: Basic Principles, Clinical Applications and Related Fields*, Williams & Wilkins, Baltimore, 1993.
- [8] R. Manuca, M. Casdagli, R. Savit, Nonstationarity in epileptic EEG and implications for neural dynamics, *Math. Biosci.* 147 (1998) 1–10.
- [9] W.S. Pritchard, D.K. Duke, K.K. Kriebel, Dimensional analysis of resting human EEG, II: surrogate-data testing indicates nonlinearity but not low dimensional chaos, *Psychophysiology* 32 (1995) 486–491.
- [10] A. Babloyantz, Destexhe, Low-dimensional chaos in an instance of epilepsy, *Proc. Natl. Acad. Sci. USA* 83 (1986) 3513–3517.
- [11] I. Dvorak, J. Siska, On some problems encountered in the estimation of the correlation dimension of the EEG, *Phys. Lett. A* 118 (1986) 63–66.
- [12] G.W. Frank, T. Lookman, M.A.H. Nerenberg, C. Essex, J. Lemieux, W. Blume, Chaotic time series analysis of epileptic seizures, *Physica D* 46 (1990) 427–438.
- [13] C.J. Stam, T.C.A.M. van Woerkon, W.S. Pritchard, Use of non-linear EEG measures to characterize EEG changes during mental activity, *Electroenceph. Clin. Neurophysiol.* 99 (1996) 214–224.
- [14] J.P. Pijn, D.N. Velis, M.J. van der Heyden, J. DeGoede, C.W.M. van Veelen, F.H. Lopes da Silva, Nonlinear dynamics of epileptic seizures on basis of intracranial EEG recordings, *Brain Topography* 9 (1997) 249–270.
- [15] M. Casdagli, Nonlinear prediction of chaotic time series, *Physica D* 35 (1989) 335–356.
- [16] P. Grassberger, I. Procaccia, Characterization of strange attractors, *Phys. Rev. Lett.* 50 (1983) 346–349.
- [17] A. Wolf, J.B. Swift, H.L. Swinney, J.A. Vastano, Determining Lyapunov exponents from a time series, *Physica D* 16 (1985) 285–317.
- [18] L. Diambra, C.P. Malta, Nonlinear models for detecting epileptic spikes, *Phys. Rev. E* 59 (1999) 929–937.
- [19] L. Diambra, A. Capurro, A. Plastino, Neural networks that learn how to detect epileptic spikes, *Phys. Lett. A* 241 (1998) 61–66.
- [20] A. Capurro, L. Diambra, D. Lorenzo, O. Macadar, M.T. Martín, C. Mostaccio, A. Plastino, J. Pérez, E. Rofman, M.E. Torres, J. Velluti, Human brain dynamics: the analysis of EEG signals with Tsallis information measure, *Physica A* 265 (1999) 235–254.
- [21] K. Lehnertz, C.E. Elger, Can epileptic seizures be predicted? Evidence from non-linear time series analysis of brain electrical activity, *Phys. Rev. Lett.* 80 (1998) 5019–5022.
- [22] J. Martinerie, C. Adam, M. Le Van Quyen, M. Baulac, S. Clemenceau, B. Renault, F.J. Varela, Epileptic seizures can be anticipated by non-linear analysis, *Nat. Med.* 4 (1998) 1173–1176.
- [23] L. Diambra, A. Plastino, Modeling time series using information theory, *Phys. Lett. A* 216 (1996) 278–282.
- [24] A.M. Albano, A.F. Mees, G.C. de Guzman, P.E. Rapp, Data requirements for reliable estimation of correlation dimension, in: H. Degn, A. Holden, L.F. Olsen (Eds.), *Chaos in Biological Systems*, Plenum, New York, 1987, pp. 207–219.
- [25] F. Lopes da Silva, EEG analysis: theory and practice, in: E. Niedermeyer, F. Lopes da Silva (Eds.), *Electroencefalography; Basic Principles, Clinical Applications and Related Fields*, Williams & Wilkins, Baltimore, 1993, pp. 1097–1123.
- [26] J. Theiler, On the evidence for low dimensional chaos in an epileptic EEG, *Phys. Lett. A* 196 (1995) 335–341.
- [27] F. Takens, Detecting strange attractors in turbulence, in: D. Rand, L.S. Young (Eds.), *Lecture Notes in Mathematics*, Vol. 898, Springer, New York, 1981, pp. 365–816.
- [28] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich, L.S. Tsimring, The analysis of observed chaotic data in physical systems, *Rev. Mod. Phys.* 65 (1993) 1331–1392.

- [29] C.E. Shannon, W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Chicago, 1949.
- [30] E.T. Jaynes, *Information theory and statistical mechanics*, *Phys. Rev.* 106 (1957) 620–630.
- [31] A. Albert, In: *Regression and Moore–Penrose Pseudoinverse*, Academic Press, New York, 1972.
- [32] A.M. Fraser, H.L. Swinney, Independent coordinates for strange attractors from mutual information, *Phys. Rev. A* 33 (1986) 1134.
- [33] F. Takens, Detecting nonlinearities in stationary time series, *Int. J. Bifurcation Chaos* 3 (1993) 241–256.
- [34] H.D.I. Abarbanel, M.B. Kennel, Local false nearest neighbors and the dynamical dimension from observed chaotic data, *Phys. Rev. E* 47 (1993) 3057–3068.
- [35] R. Manuca, R. Savit, Model misspecification test, model building and predictability in complex systems, *Physica D* 93 (1996) 78–100.
- [36] M. Green, R. Savit, Dependent-variables in broad-band continuous-time series, *Physica D* 50 (1991) 521–544.
- [37] D. Johnston, T.H. Brown, Brain slices, in: P.A. Schwartzkroin, H.V. Wheal (Eds.), *Electrophysiology of Epilepsy*, Academic Press, London, 1984, pp. 25–29.