ELSEVIER

Physica D 96 (1996) 344–354

PHYSICA D

Entropy quantification of human brain spatio-temporal dynamics ★

Laurent Pezard *, Jacques Martinerie, Johannes Müller-Gerking, Francisco J. Varela, Bernard Renault

Unité de Psychophysiologie Cognitive, LENA (CNRS URA 654-UPMC), Hôpital de la Salpêtrière, 47 Bd de l' Hôpital, 75651 Paris Cedex 13, France

Abstract

We present a procedure to quantify spatio-temporal dynamics applied here to brain surface recordings during three distinct cognitive tasks. The method uses 19 sites of EEG recording as spatial embedding for the reconstruction of trajectories, global and local linear indices, and non-linear forecasting methods to quantify the global and local information loss of the dynamics (K-entropy). We show that K-entropy can differentiate between raw and multivariate phase random surrogate data in a significant percentage of EEG segments, and that relevant non-linear indices are best studied in time segments not longer than 4 s. We also find a certain complementarity between local non-linear and linear indices for the discrimination between the three cognitive tasks. Moreover, localized projections onto electrode site of K-entropy provide a new kind of brain mapping with functional significance.

Keywords: Non-linear forecasting; Electroencephalogram; K-entropy; Spatio-temporal dynamics

1. Introduction

Some decades after the original discovery of human electroencephalogram by Hans Berger [1], the development of multichannel recording has led to an appreciation of the importance of spatio-temporal information to understand brain functioning. Originating in Antoine Rémond spatio-temporal maps [2], these studies initiated several directions of work from the description of time evolution of brain maps [3] to the study of spatio-temporal linear correlations [4]. However, dynamical systems theory has introduced new tools to quantify brain dynamics [5,6]. The methods usually applied in this area use correlation dimension computation based on time delay embedding of single electrode data. As pointed out by Albano and Rapp [7] this procedure seems to be poorly adapted to an inhomogeneous spatially extended system such as the brain. A variety of approaches making use of data from several channels have been introduced so far:

(1) Maps of correlation dimensions have been obtained from time delay embedding for each channel [8]. However, it has been shown that, in the case of spatially extended systems, the dimensions obtained from one site of observation may be erroneous and lead thus to false conclusions [9].
(2) Instead of using time delay, multiple spatial samples have been used as embedding

★ This article is dedicated to Antoine Rémond, founder of the LENA and pioneer in spatio-temporal analysis of human electroencephalogram.

* Corresponding author: Fax: +33 1 44 24 39 54; e-mail: lenalp@ext.jussieu.fr.

0167-2789/96/$15.00 Copyright © 1996 Elsevier Science B.V. All rights reserved
PII S0167-2789(96)00032-2

to obtain global dynamical indices [10,11]. This procedure considers a global system without taking into account the characteristic lengths of the system [12]. Moreover this embedding technique is restricted to sufficiently stationary spatio-temporal dynamics [13] which is in general not true for the brain.

(3) Inspired by synergetics principles [14], spatial modes of brain dynamics have been defined [15]. However the physiological or functional meaning of the principal modes are difficult to interpret, all the more that this method has been reported for only one subject. Furthermore this technique concentrates on the characterization of the dynamics associated with phase transitions and does not provide characterization of the dynamics between transitions.

We present here a method which combines multichannel embedding (i.e. spatial instead of time-lag embedding) and non-linear forecasting methods. This method quantifies the rate of loss of predictability as a function of prediction time for each localized site of recording *and* for the global or average underlying neural dynamics. In this paper we shall refer to the rate of loss of predictability as *K-entropy* (Kolmogorov entropy), for reasons that are clarified below. In order to guard against spurious results we constructed a series of surrogate data and compared the indices obtained from them with those from the original data [16,17]. Furthermore, since linear characteristics may possibly underlie the main part of the trajectories we have also calculated a global linear index computed on the basis of the eigenvalues of the observation marix, and a local linear index related to the characteristic time of the autocorrelation function for each EEG channel.

We have studied EEG signals, in segments of various lengths, obtained from nine subjects in three experimental conditions: resting with eyes closed, staring at a spotlight and a simple cognitive task which consists in an auditory target detection.

Our purpose here was to answer the following question: Is it possible to distinguish the three experimental conditions over an optimal time segment based on the statistical significance of the K-entropy index? Additionally, we introduced a brain cartography of *local* values of K-entropy and of linear measures to explore their potential neurobiological significance.

2. Numerical methods

The observation of brain activity through EEG recordings can be considered as a measurement function H of a state of activity $\xi(t) \in \mathbf{R}^{\mathrm{K}}$:

$$H: \mathbf{R}^{\mathrm{K}} \rightarrow \mathbf{R}^{k}; \qquad \mathbf{X}(t) = H(\xi(t)),$$

leading to k time series $x_j (j = 1, \ldots, k)$. For an observation period T and for each recording site j, a series of N values $x_j(i\tau)$ are obtained ($T = N\tau$, where τ denotes the sampling interval and $i = 1, \ldots, N$ is the number of samples obtained from the time origin t_0, here $t_0 = 0$). In this study we have analyzed the same EEG recording divided into an increasing number of segments of different length ($N = 8192(T), 4096(\frac{1}{2}T), 2048(\frac{1}{4}T), 1024(\frac{1}{8}T)$ and $512(\frac{1}{16}T)$ vectors). Each series x_j, has been normalized to zero mean and unit variance.

2.1. Multichannel embedding

We built the observation matrix $\mathbf{A}$ (N rows and k columns) $\mathbf{A} = \{\mathbf{X}_i\}(i = 1, \ldots, N)$ from the N values obtained from the k recording sites, that is the N k-dimensional vectors $\mathbf{X}_i$ thus constructed:

$$\mathbf{X}_i = \{x_1(i\tau), x_2(i\tau), \ldots, x_k(i\tau)\}.$$

Analogies exist between this method and the time-lag method: the inter-electrode distance corresponds to the time delay and the sampling interval acts as the jump between vectors. Thus in the case of invariants computation such as the correlation dimension, the results may depend on the electrode placement and the sampling frequency.

2.2. Singular value decomposition (SVD)

SVD of the observation matrix was computed, the axes for which the variance was inferior to the noise level were ignored with no significant loss of information [18,19]. In our case, the threshold was set to 10^{-2} which is an upper limit for the estimation of noise variance measured directly at the acquisition chain after injection of simulated signals. A k-d tree partition [20,21] was used *in the SVD reduced space* to determine the near neighbors of each point to be used in the prediction method. However, the prediction is based on the complete matrix (i.e. R^{19}).

2.3. Non-linear forecasting

Non-linear forecasting of the system's time evolution, for a prediction time p, is realized by approximation of the function $G: \mathbf{X}_{i+p} = G(\mathbf{X}_i)$. Among the several methods available [22] we chose a simple local linear model [23]. The set of data is divided into two parts of equal size. The first one is used as a learning set to build a model of the system's dynamic. The second part is a test set which is compared to the predicted series. To compute the predictor $\mathbf{X}^*_{i+p}$ for some vector $\mathbf{X}_i$ in the test set, one determines the simplex of the vector $\mathbf{X}_i$, which corresponds to the set of the $2k+1$ nearest neighbors that belong to the learning set.

The predictor $\mathbf{X}^*_{i+p}$ is a weighted sum of the successors of the simplex points, p steps ahead:

$$\mathbf{X}^*_{i+p} = \sum_{n=1}^{2k+1} w_n \cdot \mathbf{X}_{n+p},$$

where w_n is defined as:

$$w_n = \frac{d_n^{-2}}{\left(\sum_{1=1}^{2k+1} d_1^{-2}\right)} \quad \text{and} \quad \sum_{n=1}^{2k+1} w_n = 1$$

with $d_n = |\mathbf{X}_i - \mathbf{X}_n|$ (Euclidean distance).

Predictors $\mathbf{X}^*_{i+p}$ are computed for all the $\mathbf{X}_i$ belonging to the test set ($\frac{1}{2}N$ vectors). The linear correlation coefficient $\rho_j(p)$ between the jth component of the observed $\mathbf{X}_{i+p}$ (i.e. $x_j((i+p)\tau)$) and the jth component of the predicted $\mathbf{X}^*_{i+p}$ (i.e. $x^*_j((i+p)\tau)$) is computed from the series $\{x_j; x^*_j\}$. Thus for each prediction time p ($p = 1, \ldots, 10$), the vector $\mathbf{P}(p) = \{\rho_j(p)\}(j = 1, \ldots .k)$ of localized correlation coefficients is obtained. This procedure leads to a prediction curve for each channel, which we will enter in the computation of the *localized K-entropy* (K_j). The average value $\langle\rho(p)\rangle = (1/k)\sum_{j=1}^{k} \rho_j(p)(p = 1, \ldots, 10)$ constitutes the average prediction curve which will enter in the computation of the *average K-entropy* ($\langle\mathrm{K}\rangle$).

2.4. Entropy estimation

Following Wales [24], we now define the K-entropy (K), equivalent to Kolmogorov entropy, as the rate of loss of predictability in our signals from the relation:

$$\ln(1-\rho) = 2\ln(S_0/(2\sigma)) + 2\mathrm{K}p\tau,$$

where S_0 is an index of representativity of the learning set, σ the variance of the observed values, $p\tau$ the first few prediction steps. In our case, we used the first three points of each prediction curve (either localized or average) to estimate K and S_0.

2.5. Surrogate data construction and linearity test

Surrogate data are constructed from raw data by randomizing their phases in the Fourier domain [25]. To analyze the presence of non-linearities in multichannel EEG recordings, we used multivariate surrogates data that preserve the cross-correlation between channels [17]. For EEG scalp recordings, the amplitude distribution of the signals can be considered as gaussian. Therefore, our procedure does not include "gaussianization" of the data [27].

For each segment of EEG data, average K-entropy ($\langle\mathrm{K}\rangle$) was obtained from the raw data and an ensemble of surrogate data. Sets of 39 surrogate data were used to ensure a confidence level of 0.05 for the one-tailed test of the null hypothesis. In order to ensure reliable null hypothesis test we used the Monte-Carlo probability [26,27] as a robust empirical measure of significance [28]. We accept the dynamics to be non-linear when K-entropy for the surrogate data was significantly higher than that for the raw data.

A less refined procedure can be used to test whether the matched groups of subjects and of surrogate data differ for each experimental condition and each segment length. For each segment length and for each subject, a grand average ⟨K⟩ was computed over all segment, for each experimental condition. A similar calculation was performed for the surrogate values, defined as a grand average over the mean values obtained for each set of 39 surrogates. For each experimental condition and each segment length, a Wilcoxon matched pair signed ranks test was used to test the hypothesis that K-entropy for the subjects is lower than that for the surrogate values. This procedure also allowed a general indication concerning the difference of S_0 between raw and surrogate data.

2.5.1. Linear indices

To compare linear complexity [29] between experimental conditions, we used a *global* index computed from the results of SVD of the observation matrix, defined as follows:

$$\Lambda = -\left(\sum_{\nu=1}^{k} \lambda_\nu \log(\lambda_\nu)\right) \Big/ \log(k),$$

where k is the number of channels and λ_ν denotes the νth contribution to the variance ($\lambda_\nu = d_\nu^2 / \sum_{\epsilon=1}^{k} d_\epsilon^2$ with d_ν is the singular value of the νth component). Λ varies from 0 for the case of one non-zero singular value to 1 for equidistributed white noise, and it quantifies the scatter of points of the trajectory.

For each channel and in each experimental condition, the unbiased estimation of autocorrelation function,

$$f(\tau) = 1/(N-\tau) \sum_{t=1}^{N-\tau} x_t \cdot x_{t+\tau}$$

was computed. This index was computed from detrended data: the dc drift has been removed by linear regresssion [30]. The characteristic time Γ, defined as $f(\Gamma) = (1/e)$ for each channel (denoted Γ_j for channel j), constitutes a localized linear index.

3. Physiological data

The data used here had been previously analyzed with a less elaborate method [31].

3.1. Subjects

Nine young right-handed healthy adults (five females and four males) aged from 19 to 34 yrs old (mean: 24.9 ± 3.8) consituted the experimental population.

3.2. Materials

The EEG signals were recorded by 19 electrodes (Sn) placed on the scalp according to the 10–20 international electrode placement system [32]. The differences of potentials were measured between the scalp and the linked ear lobes, used as a common reference. The sample frequency was 250 Hz (digitizer 12 bits, sample and non-hold device with inter-channel sampling rate equal to 8 μs). The signal was filtered between 0.08 and 35 Hz. EEG signals were corrected for ocular artifacts using an off-line numerical method [33].

3.3. Experimental conditions

Three experimental conditions are studied below. In the first one, the subjects were asked to keep relaxed with eyes closed ("resting"). During the second condition, they were asked to stare at a red spotlight ("eyes open"). In the third one the subjects performed a binaural auditory odd-ball ("odd-ball"). This task consists in the detection of a deviant stimulus randomly presented in a series of frequent ones. The deviant stimulus (30%) were a high pitch tone (1500 Hz), and the frequent or standard (70%) a low pitch (550 Hz). The stimuli duration was 150 ms with a random interstimuli interval between 2 and 3 s. The subjects were asked to press a button with their right index finger, with both accuracy and speed, as soon as they detected the deviant stimulus.

3.4. Data structure

For each subject and each experimental condition a 32.8 s segment of EEG data free of artifact was selected (for one subject, the eyes closed condition lasted only for 14 s). These segments were analyzed for the five temporal windows defined above ($T, \frac{1}{2}T, \ldots, \frac{1}{16}T$). Thus, from each 32.8 s segment, we obtained: Two segments of 4096 vectors ($\frac{1}{2}T$), four segments of 2048 vectors ($\frac{1}{4}T$),..., and 16 segments of 512 vectors ($\frac{1}{16}T$). (For the 14 s period, we studied: One segment of 2048 vectors, three segments of 1024 vectors and seven segments of 512 vectors). In the odd-ball condition, the segments were not locked to the stimulus.

4. Results

4.1. Raw and surrogate data comparison

On the average, the correlation coefficients $\langle\rho\rangle$ are higher for the surrogate data than for the raw data, as seen in Fig. 1 which plots $\ln(1-\langle\rho\rangle)$ as a function of the prediction step p. However, the K-entropies obtained from the slope of these plots, for the different experimental conditions, yield a significant percent of segments where, as expected, the surrogates have a higher value than the raw data (see Table 1). Specifically, the percentage of segments where the null hypothesis (distinguishing surrogates and original data) could be rejected is not the same for the different segment lengths. Variations between subjects are illustrated with the results in the case of 1024 vectors ($\frac{1}{8}T$) presented in Table 2. The cases of 1024 and 512 vectors ($\frac{1}{8}T$ and $\frac{1}{16}T$) provide the most convincing evidence for non-linearity by both the examination of individual segments and the global comparison between the groups of raw and surrogate data (see Table 3).

Difference between raw and surrogate data can also be studied on the basis of the representativity of the learning set (S_0). The results of the comparisons are given in Table 4 showing that the representativity is higher for surrogate data than for raw data, specially for the shorter time segments ($\frac{1}{16}T, \frac{1}{8}T, \frac{1}{4}T$).

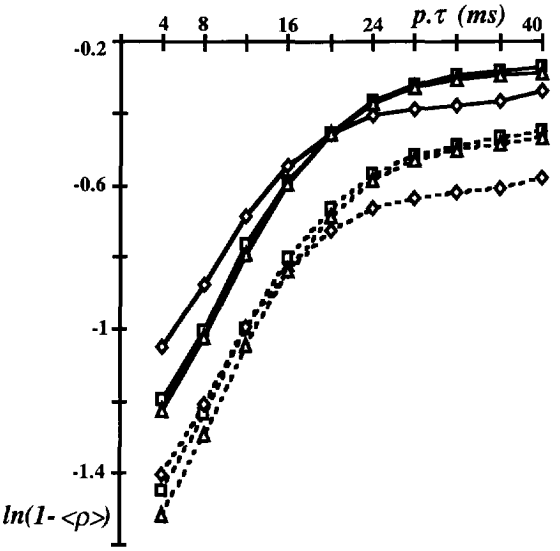


Fig. 1. Average prediction curves ($\ln(1-\langle\rho\rangle)$ as a function of prediction time) for raw (solid lines) and surrogate data (dotted lines) in the three experimental conditions (□: resting with eyes closed, △: odd ball paradigm and ◇: eyes open) for the case of $\frac{1}{8}T = 1024$ vectors.

For longer segments this trend is less clear. These differences for S_0 suggest that the surrogate data construction tends to increase the representativtiy of the learning set as compared to the test set. Nevertheless, no systematic link was noted between changes in S_0 and the significance of the K-entropy.

4.2. Comparisons between segments length

The preceding results show that multichannel EEG data can hardly be differentiated from a multivariate

Table 1
Number of segments for which the null hypothesis of linearly correlated noise can be rejected (with a significance level of 0.05 for one-tailed test) over the total number of segments for each experimental condition and each length of EEG segments. ($T = 8192$ vectors).

	Resting	Eyes open	Odd-ball
T	4/8(50%)	1/9(11%)	2/9(22%)
$T/2$	2/16(13%)	1/18(6%)	2/18(11%)
$T/4$	2/33(6%)	3/36(8%)	8/36(22%)
$T/8$	11/67(16%)	13/72(18%)	11/72(15%)
$T/16$	23/135(17%)	24/144(17%)	21/144(15%)

Table 2
Number of segments for which the null hypothesis of linearly correlated noise can be rejected at a significance level of 0.05 (one-tailed test) over the total number of segments for each experimental condition and for each subject (subj.) for the case of $\frac{1}{8}T = 1024$ vectors (4 s).

	Resting	Eyes open	Odd-ball
Subj. 3106	0/8(0%)	0/8(0%)	1/8(13%)
Subj. 3124	1/8(13%)	1/8(13%)	1/8(13%)
Subj. 3125	0/3(0%)	2/8(25%)	1/8(13%)
Subj. 3227	2/8(25%)	2/8(25%)	2/8(25%)
Subj. 3257	0/8(0%)	0/8(0%)	2/8(25%)
Subj. 3271	3/8(38%)	5/8(63%)	1/8(13%)
Subj. 3279	1/8(13%)	2/8(25%)	1/8(13%)
Subj. 3281	3/8(38%)	0/8(0%)	0/8(0%)
Subj. 3283	1/8(13%)	1/8(13%)	2/8(25%)

Table 3
Results of Wilcoxon matched pairs signed ranks tests (in bold, $P < 0.05$) for average K-entropy ($\langle K \rangle$), between the associated groups of raw and surrogate data, for each window length and each experimental conditions. Probability values are given for the one-tailed test that real values are lower than surrogates.

	Resting		Eyes open		Odd-ball	
	Σ_{rank}	P	Σ_{rank}	P	Σ_{rank}	P
T	34*	**0.012**	19	0.633	30	0.213
$T/2$	21*	0.371	19	0.633	29	0.248
$T/4$	26	0.267	20	0.590	36	0.065
$T/8$	41	**0.014**	41	**0.014**	44	**0.004**
$T/16$	43	**0.006**	44	**0.004**	42	**0.010**

* number of subjects = 8; $T = 8192$ vectors.

linearly correlated noise for long segments (T and $\frac{1}{2}T$, i.e. 32.8 and 16 s). For shorter segments ($\frac{1}{8}T$ and $\frac{1}{16}T$, i.e. 4 and 2 s), multichannel EEG data differs from noise in two ways: (1) the average K-entropy ($\langle K \rangle$) is lower for raw data than for surrogate data (Fig. 2a) and (2) the representativity of the learning set is lower for the raw data than for the surrogate data (Fig. 2b). These two results can be related to two aspects of surrogate data construction: (1) the destruction of non-linear structure causes an increase of K-entropy for surrogate data and (2) since the surrogate data are phase randomized over the whole segment, it can cause an homogenization of the whole segment leading thus to a better representativity of the learning set for surrogate data than for raw data. The increase

Table 4
Results of Wilcoxon matched pairs signed ranks tests (in bold, $P < 0.05$), for the representativity of the learning set (S_0), between the associated groups of raw and surrogate data, for each window length and each experimental condition. Probability values are given for the one-tailed test that the real values are higher than those of the surrogates.

	Resting		Eyes open		Odd-ball	
	Σ_{rank}	P	Σ_{rank}	P	Σ_{rank}	P
T	30*	0.058	40	**0.020**	35	0.082
$T/2$	27*	0.125	33	0.125	36	0.065
$T/4$	45	**0.002**	43	**0.006**	44	**0.004**
$T/8$	45	**0.002**	45	**0.002**	45	**0.002**
$T/16$	45	**0.002**	45	**0.002**	45	**0.002**

* number of subjects = 8; $T = 8192$ vectors.

of Λ as a function of the segment length confirms an increase in the scatter of points (Fig. 3).

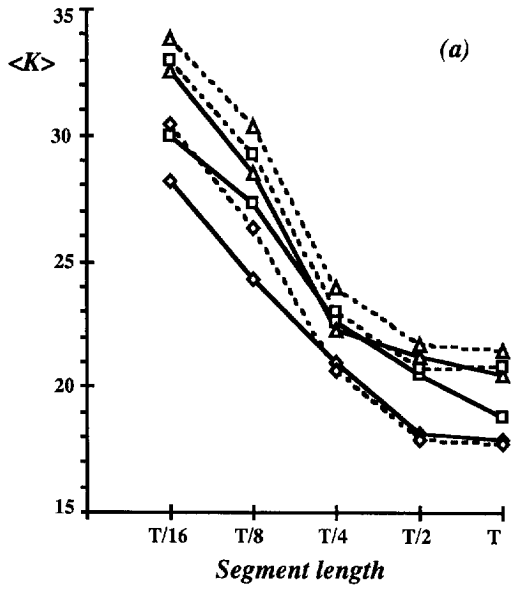
The rest of the analysis is focused on the shorter time segments ($\frac{1}{8}T$ and $\frac{1}{16}T$), since they appear to be the relevant cases to apply non-linear methods which we also compare with linear indices.

4.3. Comparisons between experimental conditions

These have been performed using average ($\langle K \rangle$) and localized K-entropy (K_j) and global (Λ) and localized (Γ_j) linear indices over all windows. Friedman two way analysis of variance by ranks was used to test the null hypothesis that the three experimental conditions have been drawn from the same population. When this null hypothesis was rejected ($P \leq 0.05$) multiple two-way comparisons between conditions were computed [34].

4.3.1. Results on global indices

The experimental conditions can be differentiated using average K-entropy only for the case of $\frac{1}{8}T = 1024$ vectors ($F_r = 6.89$, $P < 0.05$) and not for the case of $\frac{1}{16}T = 512$ vectors ($F_r = 2.67$, $P > 0.05$). Comparisons between each pair of experimental conditions shows that K-entropy is significantly lower for the eyes open condition than for the odd-ball condition ($|\Delta S_{rank}| = 11$, $P < 0.05$, two-tailed test). No difference was found for S_0 and for Λ.



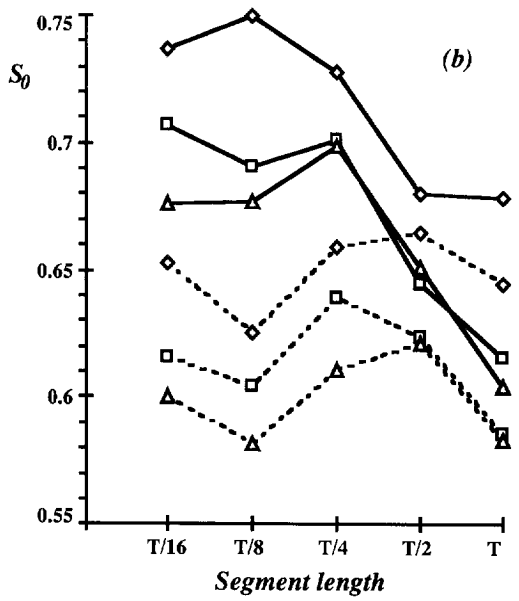


Fig. 2. Variations of the average K-entropy ($\langle K \rangle$ in s^{-1}, (a)) and of the learning set representativity (S_0, (b)) as a function of the segment length for raw (solid lines) and surrogate data (dotted lines) in the three experimental conditions (□: resting with eyes closed, △: odd-ball paradigm and ◇: eyes open).

4.3.2. Results on localized indices

Localized K-entropy (K_j) can also be used to compare experimental conditions. The maps of K_j obtained for $\frac{1}{8}T$ (1024 vectors) are shown on Fig. 4, and are similar for $\frac{1}{16}T$ (512 vectors). Friedman two-way analysis of variance by ranks was used to test the null hypothesis that the three experimental conditions have been drawn from the same population for each location. Multiple two-way comparisons between conditions were computed when the null hypothesis was rejected ($P \leq 0.05$, two-tailed test). The results are depicted in Fig. 5 (upper) for the case of $\frac{1}{8}T = 1024$ vectors (similar to the case for $\frac{1}{16}T$ except for the C_z electrode). The major differences are found for the electrodes placed on the right occipital and parietal lobes; the results are mainly linked to the gross difference between eyes open and eyes closed conditions (the occipital lobe corresponds to visual areas). It is worth noting that visual inputs are associated with a decrease of entropy in the corresponding areas.

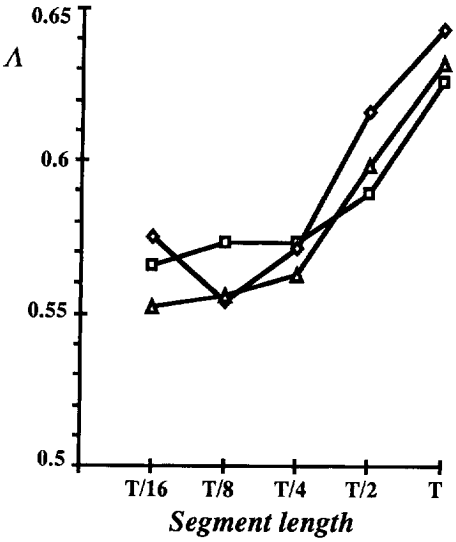


Fig. 3. Variations of Λ as a function of the segment length for original data in the three experimental conditions (□: resting with eyes closed, △: odd-ball paradigm and ◇: eyes open).

Fig. 5(b) and Table 5 display the results of the statistical tests (Friedman two-way analysis of variance and multiple two-way analysis) computed with the values of Γ_j for the case of $\frac{1}{8}T = 1024$ vectors. There are more significant electrode site differences for the case of localized K-entropy than for the autocorrelation values and shared significant sites are restricted to the right parieto–occipital lobe.

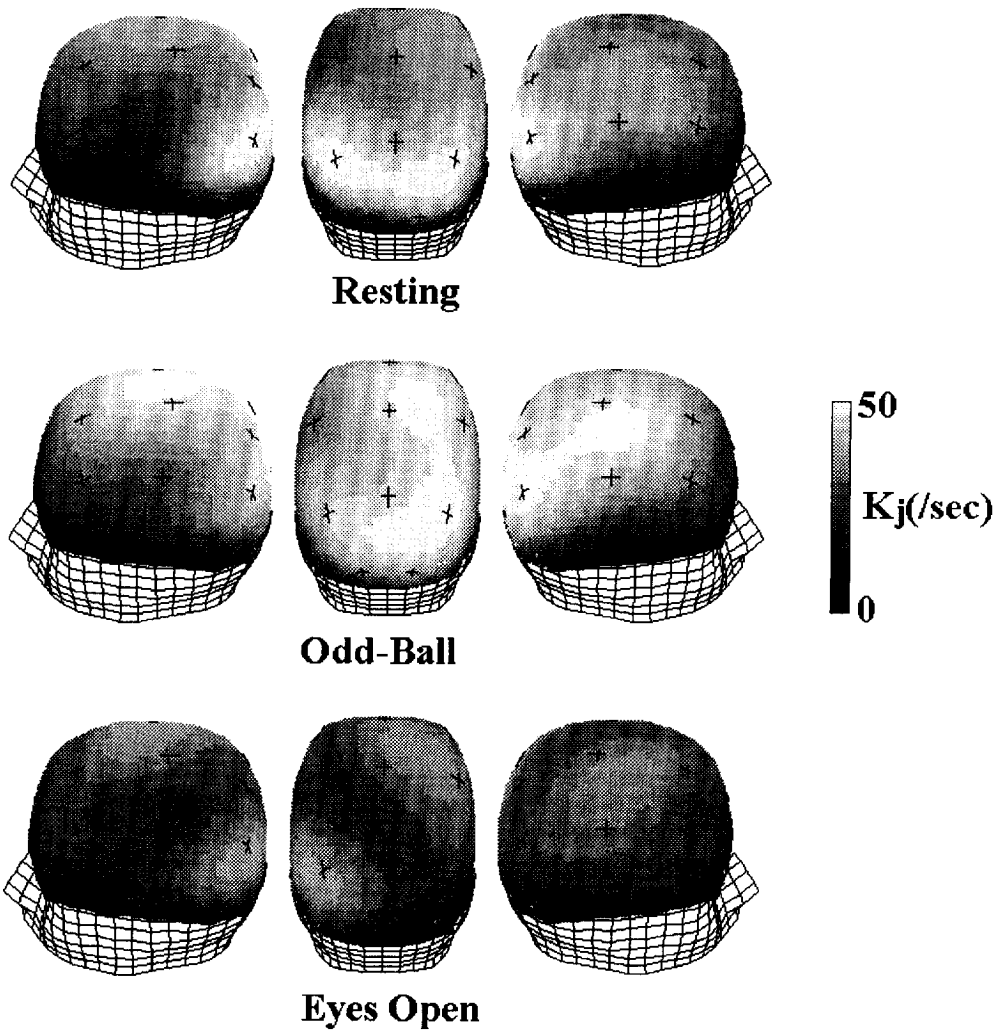


Fig. 4. Maps of local entropy (K_j) for the three experimental conditions. Different views (left, right and back side) are shown for each experimental condition. The extrapolation between local values was computed using spherical splines functions.

5. Discussion

On the whole, we have shown that the association between spatial embedding, linear and non-linear indices permits a certain characterization of the brain dynamics underlying simple cognitive tasks. We have also shown that the choice of an appropriate time segment and the use of localized indices adds significantly to EEG studies. The present method could be used in other inhomogeneous, spatially extended systems.

5.1. Non-linear methods and EEG

Previous studies using multichannel EEG analysis based on the calculation of dimension (correlation integral and Takens estimation) concluded that there was a relatively weak trace of non-linear dynamic present in scalp recordings [17,35]. The present results using K-entropy reach a similar conclusion which is not surprising given that even intracortical and single-cell analysis yield only weak

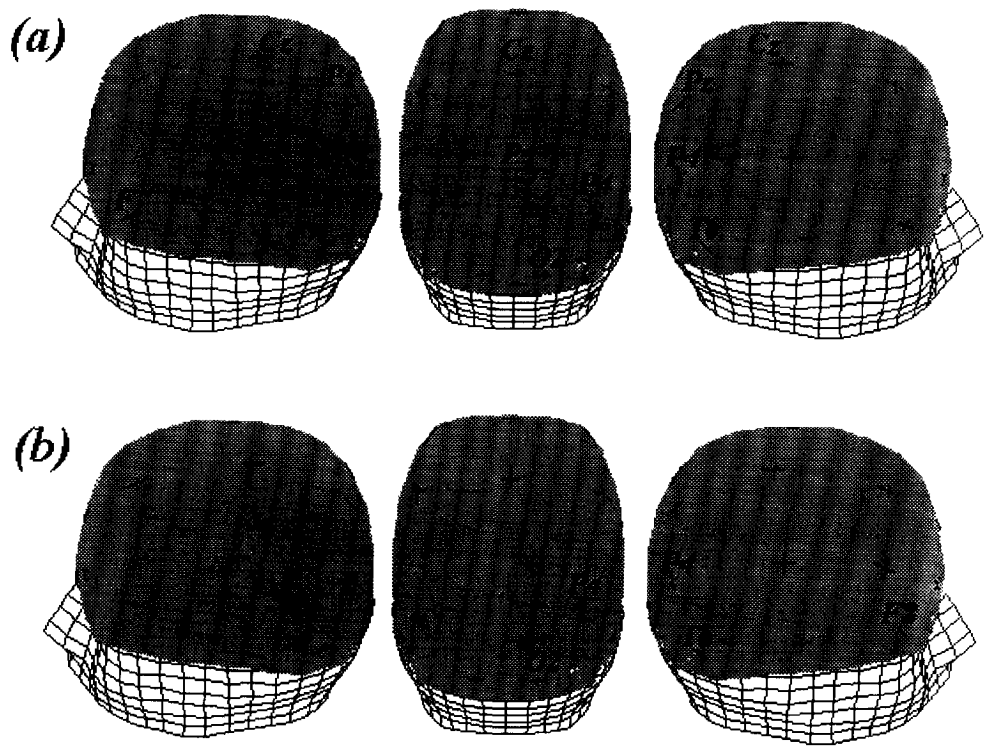


Fig. 5. Maps of significant differences between experimental conditions. The marked sites correspond to significant rejection of the null hypothesis that the three experimental conditions have been drawn from the same population at this electrode (Friedman two-way analysis of variance by ranks, two-tailed test threshold: $P \leq 0.05$). Black crosses represent the electrodes for which it was not possible to reject the null hypothesis ($P > 0.05$). (a): Results obtained for localized K-entropy (K_j). The significant differences obtained in the two-way comparisons between experimental conditions are the following: resting vs. odd-ball: F7; resting vs. eyes open: Pz, P4, T6; odd-ball vs. eyes open: Cz, O2, P4, T6. (b): Results obtained for characteristic time of the autocorrelation function (Γ_j). The significant differences obtained in the two-way comparisons between experimental conditions are the following: resting vs. odd-ball: F8; resting vs. eyes open: P4, F8; odd-ball vs. eyes open: O2, T6.

Table 5
Mean results and standard errors obtained for localized entropy (K_j) and characteristic time of autocorrelation function (Γ_j) for the electrodes that permit to differentiate significantly the three experimental conditions (in bold, $P \leq 0.05$) on the basis either of non-linear localized entropy or of characteristic time of autocorrelation function.

	K_j (in s^{-1})			Γ_j (in ms)		
	Resting	Eyes open	Odd-ball	Resting	Eyes open	Odd-ball
F7	**12.6 ± 2.7**	**15.9 ± 3.3**	**16.9 ± 2.1**	310.2 ± 31.2	252.6 ± 25.4	231.3 ± 29.12
F8	17.6 ± 3.1	22.5 ± 3.2	20.7 ± 1.5	**322.5 ± 22.2**	**203.0 ± 18.6**	**265.3 ± 25.2**
Pz	**42.0 ± 2.5**	**32.5 ± 3.5**	**43.7 ± 2.8**	103.0 ± 29.0	162.0 ± 20.4	85.0 ± 21.5
P4	**46.2 ± 2.8**	**31.7 ± 4.4**	**45.8 ± 4.9**	**84.0 ± 23.2**	**167.7 ± 25.0**	**70.7 ± 20.5**
T6	**26.7 ± 4.8**	**17.6 ± 4.8**	**29.2 ± 5.4**	**122.7 ± 26.6**	**191.1 ± 15.9**	**88.2 ± 18.4**
Cz	**38.0 ± 2.6**	**34.9 ± 2.8**	**44.8 ± 2.9**	149.7 ± 40.9	156.2 ± 18.2	125.6 ± 24.4
O2	**36.5 ± 3.9**	**23.4 ± 6.4**	**39.0 ± 4.7**	**97.44 ± 29.0**	**186.0 ± 20.1**	**82.6 ± 29.0**

traces of non-linear determinism with current methods [36]. The relative weakness of these traces of non-linearities possibly stems from the fact that electrical field suffers passing through the head bone and tissues, eliminating many high frequencies that play a key role in neural dynamics. Similarly this underscores the need for recording EEG signals destined for dynamical analysis with a minimum of active filtering, which is contrary to common practice. Moreover, in a spatially distributed system, it is difficult to differentiate locally between deterministic processes and noise [37]. Thus, in the case of EEG, non-rejection of the null hypothesis could be due to artifactual correlation between local neuronal activities caused by the diffusion of the electrical fields through the tissues. Other techniques such as magnetoencephalography that are not sensible to either filtration through the tissues or to field diffusion could thus be important in further studies.

Although the use of spatial embedding seems very attractive in the case of spatially distributed systems, the problem of the dimensionality of the embedding space is still an open question. It is probable that the 19-dimensional space used here is still too low for the study of more refined invariants of brain dynamics. More channels and/or spatio-temporal embedding may be used to increase the space dimension but little is known about the practical aspects of such a technique. Nevertheless the present method permits to obtain *localized* characteristics, without the drawbacks of reconstructions based on single channels.

We have found that segments lasting up to 4 s ($\frac{1}{8}T$) are an upper bound for detecting the presence of non-linearities in multichannel EEG. This duration is likely to be linked to the known intrinsic rhythms in the brain, showing a "relaxation" time for many cognitive activities of the order of 1 s (see e.g. [38]). Thus longer periods are likely to accumulate significant non-stationarities, causing the traces of non-linearity to be drowned. The 4 s segment (1024 vectors) seems to be a good compromise between this intrinsic brain rhythm and the methodological needs related to the number of the points for computation.

5.2. *Implications for neuroscience*

The spatial embedding allows to compute local indices within the context of the global dynamics. In the case of the brain, this method is meaningful since long range cooperation are anatomically present and play an important role in brain functioning [39]. Local entropy maps show specific activations related to neuronal activity. In particular we observed an increase in K-entropy over the occipital pole during the eyes closed conditions (resting and odd-ball) which are known to lead to a substantial increase in the 8–12 Hz alpha range normally considered to be a stable and predictable condition [40]. In other words, alpha increase might be an indication of increased complexity and not the converse. This method could thus permit to deal with spatial inhomogeneities such as the functional specializations and therefore become a complement to other cerebral imaging techniques. Moreover, higher sampling frequencies could help to characterize shorter EEG segments (in the order of 1 s) which are of interest for the study of cognitive functions.

It should be noted that although we have found a certain degree of complementarity between local nonlinear and local linear indices, these two differ substantially. In fact the former is calculated on the basis of the entire trajectory of the observed electrode set, while the latter is strictly independent of spatial localization. Nevertheless it seems that both measures can contribute to characterize or discriminate different aspects of cognitive dynamics.

Distinctive characterizations could not be obtained amongst all the experimental conditions (e.g. resting and odd-ball). In order to study differences between conditions within one group, the characterization should certainly be more precise (stimulus locked for example) and the experimental paradigm designed to increase the specificity of neuronal activation. The present method could also be of clinical importance to better understand diseases when considered as dysfunctional dynamics. We have reported elsewhere that average prediction curves and entropy permit to differentiate between control and pathological state (major depression) [41,42].

Our use of non-linear forecasting is mainly a pragmatic compromise between theoretical tools and some practical needs. This should not let us loose sight of the needed developments in theory leading for example a true space–time matrix reconstruction. Such advances are likely to have further impact for neurosciences.

Acknowledgements

The authors wish to thank Jean-Claude Bourzeix for his technical assistance and Jean-Louis Nandrino and an anonymous referee for their helpful comments on a previous version of this manuscript.

References

[1] H. Berger, Archiv für Psychiatry und Nervenkrankheiten 87 (1929) 527; Reprinted in P. Gloor, Electroenceph. Clin. Neurophysiol. Suppl. 28 (1969) 37.
[2] A. Rémond, Rev. Neurol. 93 (1955) 399.
[3] D. Lehmann, in: Methods of Analysis of Brain Electrical and Magnetic Signals, eds. A.S. Gevins and A. Rémond, EEG Handbook (revised series, vol. 1) (Elsevier, Amsterdam, 1987) p. 309.
[4] A. Gevins and B. Cutillo, Electroenceph. Clin. Neurophysiol. 87 (1993) 128.
[5] A. Babloyantz, M. Salazar and C. Nicolis, Phys. Lett. A 111 (1985) 152.
[6] P.E. Rapp, T.R. Bashore, J.M. Martinerie, A.M. Albano, I.D. Zimmerman and A.I. Mees, Brain Topog. 2 (1989) 99.
[7] A.M. Albano and P.E. Rapp, in: Nonlinear Dynamical Analysis of the EEG, eds. Jansen BH, Brandt ME (World Scientific, Singapore, 1993) 117.
[8] W.S. Pritchard and D.W. Duke, Psychophysiol. 29 (1992) 182.
[9] E.N. Lorenz, Nature 353 (1991) 241.
[10] A. Destexhe, J.A. Sepulchre and A. Babloyantz, Phys. Letts A 132 (1988) 101.
[11] I. Dvorak, Phys. Lett. A 151 (1990) 225.
[12] M.C. Cross and P.C. Hohenberg, Rev. Mod. Phys. 65 (1993) 851.
[13] H.D.I. Abarbanel, R. Brown, J.J. Sidorowich and L.S. Tsimring, Rev. Mod. Phys. 65 (1993) 1331.
[14] H. Haken, Advanced Synergetics (Springer, Berlin, 1987).
[15] A. Fuchs, J.A.S. Kelso and H. Haken, Int. J. Bifurc. Chaos 2 (1992) 917.
[16] P.E. Rapp, A.M. Albano, T.I. Schmah and L.A. Farwell, Phys. Rev. E 47 (1993) 2289.
[17] D. Prichard and J. Theiler, Phys. Rev. Lett. 73 (1994) 951.
[18] D.S. Broomhead and G.P. King, Physica D 20 (1986) 217.
[19] A.M. Albano, J. Muench, C. Schwartz, A.I. Mees and P.E. Rapp, Phys. Rev. A 38 (1988) 3017.
[20] J.H. Bentley, Commun. ACM 18 (1975) 509.
[21] J.H. Bentley, IEEE Trans. Softw. Engrg. 5 (1979) 333.
[22] M. Casdagli, Physica D 35 (1989) 335.
[23] G. Sugihara and R.M. May, Nature 344 (1990) 734.
[24] D.J. Wales, Nature 350 (1991) 485.
[25] J. Theiler, S. Eubank, A. Longtin, B. Galdrikian and J.D. Farmer, Physica D 58 (1992) 77.
[26] G.A. Barnard, J.R. Stat. Soc. B25 (1963) 255.
[27] A.C.A. Hope, J.R. Stat. Soc. B30 (1968) 582.
[28] P.E. Rapp, A.M. Albano, I.D. Zimmerman and M.A. Jiménez-Montaño, Phys. Lett. A 192 (1994) 27.
[29] S.D. Morgera, IEEE Trans. Systems Man Cybern. 15 (1985) 608.
[30] J. Theiler and P.E. Rapp, Electroenceph. Clin. Neurophysiol. 98 (1996) 213.
[31] L. Pezard, J. Martinerie, F. Breton, J.-C. Bourzeix and B. Renault, Electroenceph. Clin. Neurophysiol. 91 (1994) 383.
[32] H. Jasper, Electroenceph. Clin. Neurophysiol. 10 (1958) 371.
[33] G. Gratton, M.G.H. Coles and E. Donchin, Electroenceph. Clin. Neurophysiol. 55 (1983) 468.
[34] S. Siegel and N.J. Castellan, JR, Nonparametric Statistics for the Behavioral Sciences (McGraw-Hill, New York, 1988).
[35] S.A.R.B. Rombouts, R.W.N. Keunen and C.J. Stam, Phys. Lett. A 202 (1995) 352.
[36] J. Müller-Gerking, J. Martinerie, S. Neuenschwander, L. Pezard, B. Renault and F.J. Varela, Physica D 94 (1996) 65.
[37] G. Mayer-Kress and K. Kaneko, J. Stat. Phys. 54 (1989) 1489.
[38] F.J. Varela, A. Toro, E. Schwartz and E.R. John, Neuropsychologia 19 (1981) 675.
[39] M.J. Farah, Behav. Brain Sc. 17 (1994) 43.
[40] G. Dumermuth and L. Molinari, in: Methods of Analysis of Brain Electrical and Magnetic Signals, eds. A.S. Gevins and A. Rémond, EEG Handbook (revised series, vol. 1) (Elsevier, Amsterdam, 1987) p. 85.
[41] J.-L. Nandrino, L. Pezard, J. Martinerie, F. El Massioui, B. Renault, B. Jouvent, J.-F. Allilaire and D. Widlöcher, NeuroReport 5 (1994) 528.
[42] L. Pezard, J.-L. Nandrino, B. Renault, F. El Massioui, J.-F. Allilaire, J. Müller, F.J. Varela and J. Martinerie, Biol. Psychiatry, in press.