REGULAR PAPER

J. Physiol. Sci. Vol. 58, No. 4; Aug. 2008; pp. 239–247
Online Jun. 28, 2008; **doi:10.2170/physiolsci.RP004708**

A Comparison Analysis of Embedding Dimensions between Normal and Epileptic EEG Time Series

Ye YUAN[1], Yue LI[1], and Danilo P. MANDIC[2]

[1]Department of Information and Engineering, Jilin University, Changchun, China; and [2]Department of Electrical and Electronic Engineering, Imperial College, London, UK

Abstract: The embedding dimensions of normal and epileptic electroencephalogram (EEG) time series are analyzed by two different methods, Cao's method and differential entropy method. The results of the two methods indicate consistently that the embedding dimensions of EEG signals during seizure will change and become different from that of normal EEG signals, and the embedding dimensions will vary intensively during seizure, whereas the embedding dimensions of normal EEG signals basically maintains stability. The embedding dimension results also reflect the variation of freedom degree of the human brain nonlinear dynamic system (NDS) during seizure. And based on the results of Cao's method, it is also found that normal EEG signals are of some degree of randomness, whereas epileptic EEG signals have determinism.

Key words: electroencephalogram (EEG), embedding dimension(s), Cao's method, differential entropy, epilepsy.

The electroencephalogram (EEG) is the most often used signal for the analysis of epileptic seizure activity of the brain. Epilepsy, from which approximately 1% of the people in the world suffer, is a group of brain disorders characterized by the recurrent paroxysmal electrical discharges of the cerebral cortex that result in irregular disturbances of brain functions, which are associated with the significant changes of the EEG signals [1, 2]. The neuronal network composing the brain is spatially extended, and thus EEG signals capture spatial local electrical activity in the brain [3]. It is known that biological neurons can be modeled by a set of nonlinear differential equations. The minimal embedding dimensions provide the upper number of nonlinear dynamic system (NDS) freedom degrees and the minimal number of differential equations demanded for a mathematical modeling of NDS [4]. Therefore the changes of the brain NDS structure during seizure can be shown by the changes of the embedding dimensions of EEG signals if the human brain is considered as a nonlinear dynamic system. The study of EEG signals has drawn more and more attention [5–8]. In the past, many works have been done to analyze epileptic EEG signals based on nonlinear dynamic methods [9–14]. For example, Xiaoli Li *et al.* (2006) proposed an epileptic seizure prediction based on mutual information and Cao's method [9], and L. Diambra *et al.* (1999) presented a technique for the automatic detection of epileptic spikes based on an optimal embedding dimension [10]. All these methods can clearly show the difference between normal and epileptic EEG signals based on nonlinearity or statistical information; nevertheless, they do not show the changes of brain NDS structure during seizure.

Because the EEG signals vary over time, dynamic measures should be computed within a certain time scale, for which the local stationarity assumption is valid. If statistics are used that take too long to quantify the EEG, the local information will be washed away, and most of the temporal characteristics will be deprived of their diagnostic values.

In this paper, we will investigate the difference of the embedding dimensions between normal and epileptic EEG signals by using two methods, Cao's and differential entropy [15, 16]. The reason why two kinds of methods are applied is because until now there has been no criterion for determining the exact value of the embedding dimensions of a real-world time series. We hope to compare the results from two different methods for more accuracy to avoid any influence of either method.

EEG TIME SERIES

The raw EEG time series used in this paper were recorded by a Nihon Kohden EEG recorder (model 7310; Nihon Kohden Inc., Tokyo, Japan). The electrodes were placed on the scalp to record 16 channels of EEG data according to the international standard 10–20 system. The data were digitized with a 12-bit analog-to-digital converter at

Received on Mar 24, 2008; accepted on Jun 25, 2008; released online on Jun 28, 2008; doi:10.2170/physiolsci.RP004708
Correspondence should be addressed to: Yue Li, Department of Information and Engineering, 71#, Nanhu Campus, Jilin University, Changchun, 130012, China. Tel: +86-431-8515-2020, E-mail: liyue84@gmail.com

a sampling rate of 200 Hz and stored on a computer hard disk. We recorded 80-second-long EEG data, thus 16,000 sample points for each channel. Two groups of EEG time series were obtained, both of which were collected from the same 39-year-old male patient. The volunteer was relaxed in an awake state with eyes closed. One group was recorded when the patient was normal, and the other group was recorded when the patient was in an induced epileptic seizure. Some of the EEG signals used in this paper are shown in Fig. 1, (a) and (b). From Fig. 1 (b) it can be observed that the epileptic EEG time series show epileptic characteristics, such as spike discharge and slow waves.

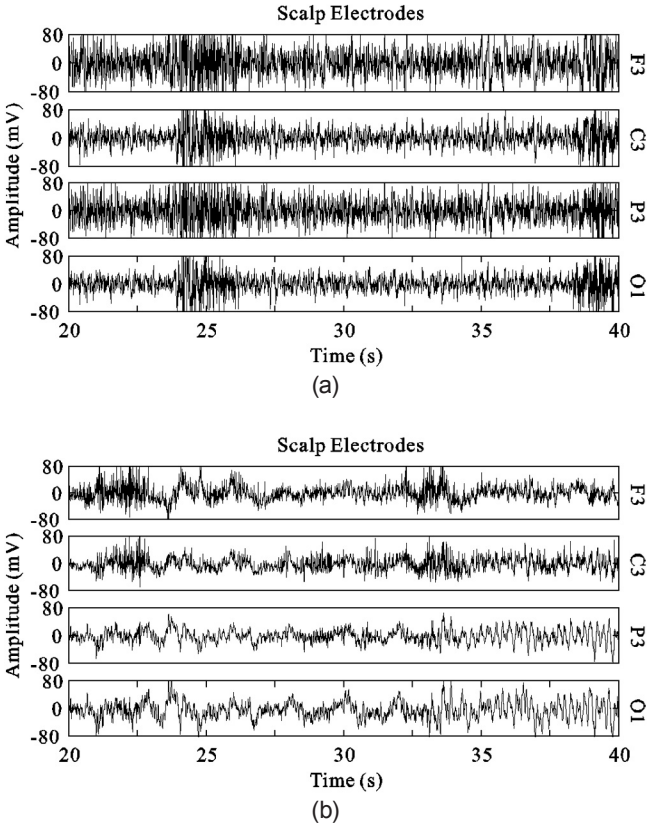


Fig.1. EEG signal examples: (**a**) Normal EEG; (**b**) Epileptic EEG.

METHODS

Cao's method. To determine the embedding dimensions of the time series $x_1, x_2, \cdots, x_N$ by Cao's method, first compute

$$E_1(d) = E(d+1)/E(d) \quad (1)$$

and

$$E_2(d) = E^*(d+1)/E^*(d) \quad (2)$$

second, plot $E_1(d)$ and $E_2(d)$ versus dimension d, and find the dimension d_0 where $E_1(d)$ stops changing from the plot. Then d_0 is the minimum embedding dimension we look for. Here, $E_2(d)$ is used to determine whether a time series is random. The steps and explanation in detail of Cao's method are shown in Appendix 1. The delay time τ, which is needed for the computation of the embedding dimension by Cao's method, is determined by the mutual information method [17].

Differential entropy method. A method based on differential entropy for determining the optimal embedding parameters is introduced in Ref. 16, which employs a single criterion — the "entropy ratio" between the phase space representation of a signal and the ensemble of its surrogates. The advantage of the differential entropy method is that it determines the embedding dimension m and delay time τ together, whereas other methods, such as one used in this paper — Cao's method — need to determine the embedding dimension after the delay time is determined. If the delay time were too small to cover the minimal time span needed to capture the dynamics of a signal, the tap input length-embedding dimension would become rather large, resulting in an increased complexity of training. However, if the delay time is greater than optimal, the nature of the resulting model becomes too discrete, resulting in failure of the filter to capture the underlying signal dynamics. Thus there is a need for an optimization method to jointly determine the embedding dimension and delay time [16].

The entropy ratio (ER) introduced in Ref. 16 is this:

$$R_{ent}(m,\tau) = I(m,\tau)(1+\frac{m\ln N}{N}) \quad (3)$$

Plot $R_{ent}(m, \tau)$ versus m and τ, then the minimum of the plot of the entropy ratio yields the optimal set of embedding parameters, namely, optimal embedding dimension m and delay time τ. The detailed steps and an explanation of the differential entropy method are shown in Appendix 2.

SIMULATIONS AND RESULTS

How the simulations are organized. If the dynamic system is invariant, then the determination of embedding dimension can be made on any segment of the measured signal. This is not true in a time-varying environment. Because the EEG time series vary over time, they must be divided into several small segments for computing embedding dimension as time evolves. The data we used here are 16-channel normal and epileptic EEG signals; the length of the signal is 80 s for each channel; the sampling frequency is 200 Hz; and therefore there are 16,000 points for each channel signal. Each channel signal is divided into 16 segments, namely S_1, S_2, ..., S_{16}; each segment S_i (i = 1,2,...,16) is of 1,000 points. Thus we can obtain 16 embedding dimension values for each channel signal. As for the 16 channels of normal or epileptic EEG time series, we can get a 16×16 matrix built by the embedding dimension values.

Results of Cao's method. For this paper, because one

plot will be produced corresponding to every EEG segment by Cao's method to determine the embedding dimension, we cannot list all the plots here. For example, we take the channel corresponding to scalp electrode F3 from 16-channel normal and epileptic EEG data, and the embedding dimension plots produced by Cao's method are shown in Fig. 2, in which the plots in the left side correspond to normal EEG segments, and the plots in the right side correspond to epileptic EEG segments. It can be observed from the first plot in the left side in Fig. 2 that the E_1 curve is in a rising process before $d = 10$ and stops changing after $d = 10$ (i.e., after $d = 10$, the E_1 curve becomes a straight line). Therefore the corresponding embedding dimension of this segment is 10, and in a similar way the embedding dimension of the EEG segment corresponding to the first plot in the right side in Fig. 2 is 16.

Cao's method is applied to compute the embedding dimensions of S_i (i = 1,2,...,16). The results of embedding dimensions for the normal and epileptic EEG time series obtained by Cao's method are shown in Table 1, in which the upper number for each cell corresponds to the embedding dimension value of the epileptic EEG segment, and the lower number corresponds to the embedding dimension value of the normal EEG segment. Let E_{ij} (i, j = 1,2,...,16) represent the value of the point of the matrix built by the embedding dimension values, whose coordinates are (i, j). The average values are computed for each row and line of the matrix. Let AV_i represent the average value of the i row of the matrix, namely, $AV_i = \frac{\sum_{j=1}^{16} E_{ij}}{16}$; let AV_j represent the average value of the j line of the matrix, namely, $AV_j = \frac{\sum_{i=1}^{16} E_{ij}}{16}$; the results of AV_i and AV_j are shown in Table 1. Apparently the value of AV_i represents the embedding dimensions of each channel of the 16-channel EEG signals, and the value of AV_j represents the embedding dimensions of the 16-channel EEG signals at some corresponding time if the 16-channel EEG signals are regarded as a whole. The variances are also computed for each row and line of the matrix; let Var_i and Var_j represent the variance of the i row and j line of the matrix, respectively. As shown in Table 1, the variances of embedding dimensions of the epileptic EEG time series are much larger than those of the normal EEG time series, which means that the embedding dimensions of the EEG time series vary more briskly during seizure than that of the normal EEG time series. The average value of the embedding dimensions (*ED*'s) of the EEG signals can be written as

$$ED = \frac{\sum_{i=1}^{16}\sum_{j=1}^{16} E_{ij}}{16 \times 16}.$$

Therefore it can be found that the average value of the embedding dimensions of normal EEG signals is ED_{normal} = 8, and the average value of embedding dimensions of epileptic EEG signals $ED_{epileptic}$ = 17; thus the average value of the embedding dimensions of epileptic EEG signals is more than 2 times larger than that of normal EEG signals. The RMSE (root mean squared error) is computed for the results of the embedding dimensions of normal and epileptic EEG signals; the results are $RMSE_{normal}$ = 0.1753 and $RMSE_{epileptic}$ = 0.3091, which also reflect that the embedding dimensions of epileptic EEG signals varies more intensely during seizure than the embedding dimensions of normal EEG signals.

Figure 3 is plotted based on the results of the embedding dimensions shown in Table 1. As shown in Fig. 3, the curves constructed by the embedding dimension values of normal and epileptic EEG time series form two fluctuation belts, but the one corresponding to the embedding dimensions of the epileptic EEG time series is much wider than the one corresponding to those of the normal EEG time series, which implies that the embedding dimensions of epileptic EEG signals have a bigger variation interval than normal EEG signals. The AV_j curves (the curve marked by black points) corresponding to normal and epileptic both fluctuate around the *ED* lines, but the AV_j curve corresponding to epileptic EEG signals fluctuates more briskly than the one corresponding to normal EEG signals, which also means that the embedding dimensions of epileptic EEG signals vary more intensely during seizure than the ones of normal EEG signals.

Results of differential entropy method. An example about how to determine the optimal embedding parameters of an EEG segment is shown in Fig. 4. Because the lowest point of the plot, whose X axis is embedding dimension m and Y axis is delay time τ, appears at (2, 7), the optimal embedding parameters of this EEG segment can be determined as $m_{opt} = 2$ and $\tau_{opt} = 7$.

In a way similar to Cao's method, the differential entropy method is applied to compute the embedding dimensions of EEG segments S_i (i = 1,2,...,16). The results of the embedding dimensions determined by the differential entropy method are shown in Table 2, and Fig. 5 is plotted based on the results. Figure 5 shows the same rule with Fig. 3, which is plotted based on the results of Cao's method, namely, that the embedding dimensions of epileptic EEG signals have a wider variation belt than those of normal EEG signals. The embedding dimensions of epileptic EEG signals vary considerably from 2 to 5, whereas the embedding dimensions of normal EEG signals basically remain constant. In Table 2, AV_i and AV_j are constant for the normal EEG signals, whose definitions are the same as those defined in *Results of Cao's method* of this paper. AV_i and AV_j vary from 2 to 3 for epileptic EEG signals. Var_i and Var_j are also computed for the results of embedding dimensions shown in Table 2. The RMSE results obtained based on the results obtained by the differential

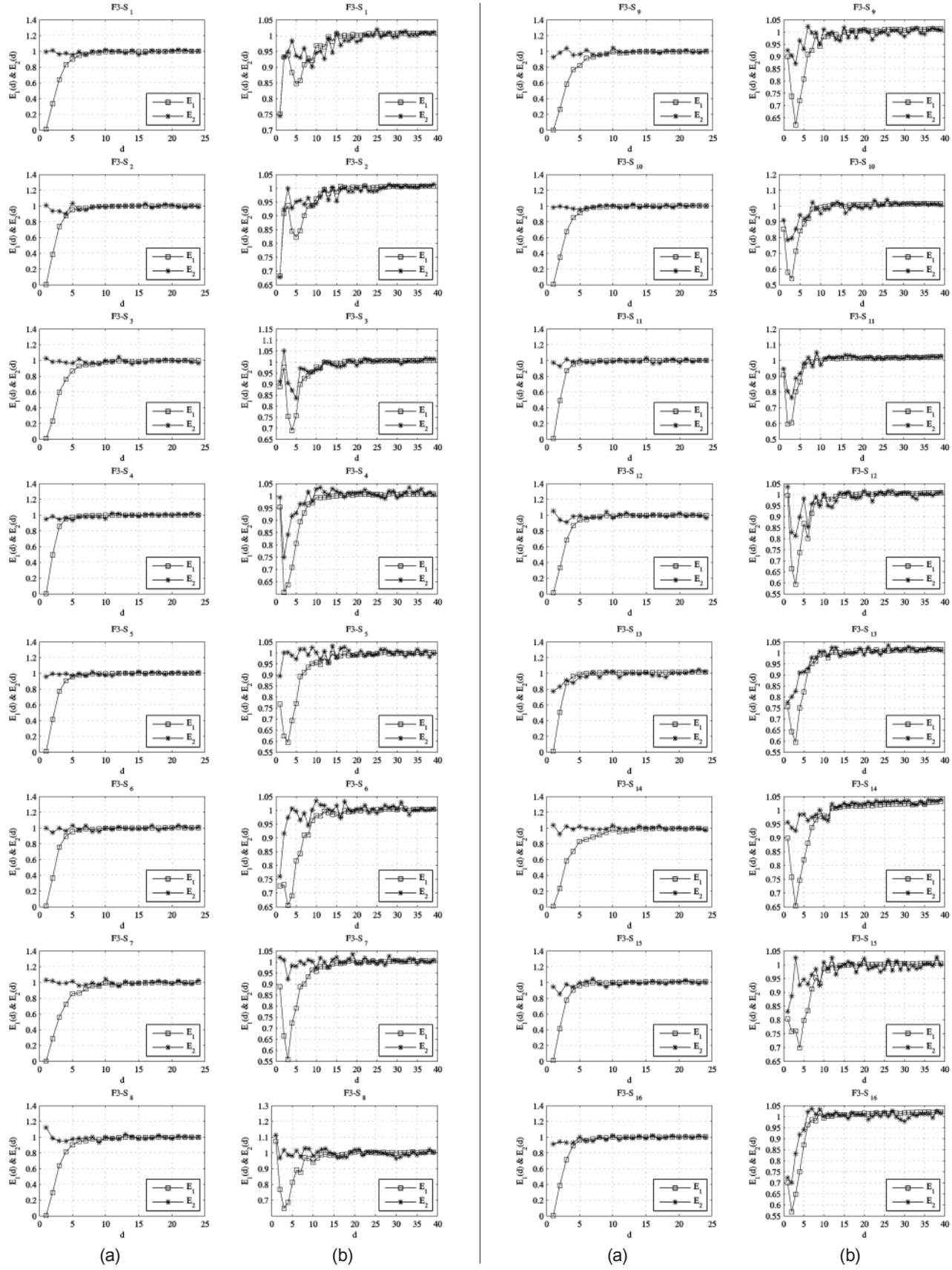
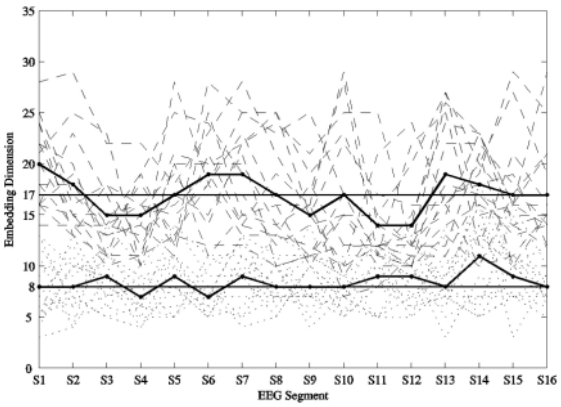


Fig. 2. Embedding dimension results of EEG segments corresponding to scalp electrode F3 by Cao's method: (**a**) Normal EEG; (**b**) Epileptic EEG. (Note: The title of each plot F3-S_i (i = 1,2,3,…,16) means the i EEG segment corresponding to scalp electrode F3.)

Table 1. Embedding dimension results obtained by Cao's method.

SE	Embedding dimensions																AV_i	Var_i
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}		
Fp1	22	20	11	15	16	19	14	13	12	10	12	10	17	18	19	12	15	14
	6	7	11	8	5	10	12	10	4	7	5	11	9	13	10	7	8	7
F3	16	23	17	10	28	18	25	23	25	17	11	22	25	12	21	16	19	30
	10	9	12	6	6	7	13	11	11	9	5	7	6	14	7	7	9	7
C3	22	15	13	14	21	20	22	25	15	10	21	24	16	19	29	25	19	26
	5	8	8	6	5	11	11	7	7	8	10	6	7	7	8	8	8	3
P3	20	15	13	18	13	18	23	16	10	15	15	12	12	21	10	15	15	14
	5	13	7	10	11	8	4	5	10	7	9	14	7	5	10	7	8	8
O1	16	13	18	16	15	18	13	13	18	12	12	11	11	17	13	14	14	6
	5	10	9	7	8	5	6	10	7	6	5	5	14	12	11	9	8	7
F7	25	17	14	19	25	24	19	17	17	28	11	15	24	23	16	14	19	24
	12	7	5	6	11	6	5	10	5	8	8	7	11	5	7	9	8	5
T3	20	15	15	16	19	14	22	24	21	25	25	12	26	23	16	21	20	19
	7	6	7	8	8	6	9	5	10	5	8	9	10	14	11	7	8	5
T5	24	16	18	12	17	15	21	10	11	12	12	13	15	23	14	12	15	18
	5	7	5	4	8	5	10	7	7	9	8	6	7	5	7	5	7	3
Fp2	24	13	11	19	12	21	14	19	16	7	8	11	14	16	14	15	15	20
	4	10	9	4	9	7	13	6	10	7	14	7	9	16	10	9	9	10
F4	28	29	22	22	18	23	28	21	15	22	17	16	16	20	27	19	21	21
	13	6	11	9	10	7	7	7	6	7	5	9	9	16	3	9	8	10
C4	16	19	14	16	12	28	25	25	19	23	15	20	27	20	16	16	19	23
	9	12	7	8	11	5	11	8	7	05	8	13	7	13	10	10	9	6
P4	14	14	10	11	20	20	20	14	16	9	14	12	16	10	17	21	15	15
	7	11	6	5	10	5	11	8	11	10	12	8	7	18	10	12	9	10
O2	18	14	14	13	17	17	17	12	14	15	12	12	18	17	15	15	15	5
	7	5	6	9	7	8	5	7	14	5	11	10	5	7	7	6	7	6
F8	18	18	13	14	11	25	15	14	11	17	11	10	22	22	18	14	16	19
	11	04	09	06	14	11	09	11	07	07	10	05	08	06	15	6	9	10
T4	21	25	23	10	20	11	15	16	16	29	14	14	27	15	16	29	19	38
	3	4	14	10	14	9	9	13	11	9	12	10	3	9	7	8	9	12
T6	16	15	11	11	13	12	12	10	11	15	10	10	13	10	11	15	12	4
	13	10	10	5	5	8	8	7	7	11	11	9	13	10	16	10	10	9
AV_j	20	18	15	15	17	19	19	17	15	17	14	14	19	18	17	17		
	8	8	9	7	9	7	9	8	8	8	9	9	8	11	9	8		
Var_j	15	21	14	12	23	21	24	27	16	47	18	19	31	19	26	22		
	10	7	6	3	8	4	8	5	7	3	8	7	7	19	10	3		

Note: SE is short for scalp electrode. For each cell of the table, the upper number corresponds to epileptic EEG time series, and the lower number corresponds to normal EEG time series. AV_i stands for the average value of the *i* row embedding dimensions, and AV_j stands for the average value of the *j* line embedding dimensions. Var_i stands for the variance of the *i* row embedding dimensions, and Var_j stands for the variance of the *j* line embedding dimensions.



The 16 dash curves are plotted by the embedding dimension values of 16-channel epileptic EEG signals. The upper curve marked by black points and the straight line are plotted by AV_j (shown in Table 1) and ED values of epileptic EEG time series.

The 16 dot curves are plotted by the embedding dimension values of 16-channel normal EEG signals. The lower curve marked by black points and the straight line are plotted by AV_j (shown in Table 1) and ED values of normal EEG time series.

Fig. 3. Scheme of variation over time of EEG's embedding dimensions obtained by Cao's method.

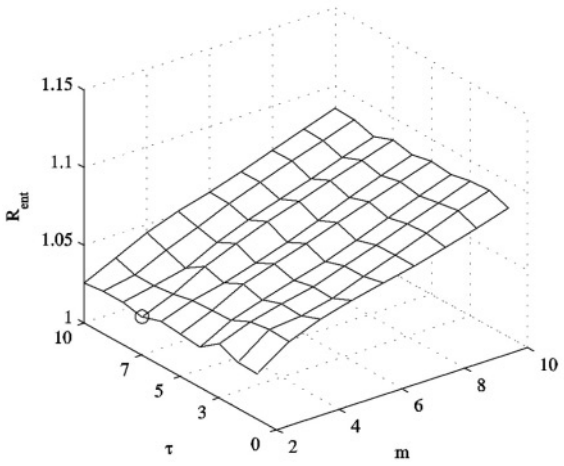


Fig. 4. An example for determining the optimal embedding parameters by the differential entropy method. The minimum of the plot indicated by an open circle yields $m_{opt} = 2$ and $\tau_{opt} = 7$.

Table 2. Embedding dimension results obtained by differential entropy method.

SE	Embedding dimensions																AV_i	Var_i
	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8	S_9	S_{10}	S_{11}	S_{12}	S_{13}	S_{14}	S_{15}	S_{16}		
Fp1	2	2	3	2	2	2	2	2	3	3	3	4	3	3	3	3	3	0.4
	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.1
F3	2	2	2	5	3	4	4	4	4	4	4	4	5	2	4	2	3	1
	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	0.1
C3	2	2	2	2	4	4	4	4	4	4	4	4	4	3	4	2	3	0.9
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
P3	2	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	0.1
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
O1	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	0.1
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
F7	2	2	2	4	3	3	3	3	3	4	4	4	4	2	5	2	3	0.9
	2	2	3	2	2	2	2	3	2	2	2	2	2	2	2	2	2	0.1
T3	2	2	2	4	4	4	3	3	4	3	3	4	3	3	4	4	3	0.6
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	3	2	0.1
T5	2	2	2	2	2	2	2	2	2	2	2	2	4	2	2	3	2	0.3
	2	2	2	2	2	2	2	2	2	2	2	3	2	2	2	2	2	0.1
Fp2	2	2	3	2	2	2	2	2	2	4	3	3	4	3	2	3	3	0.5
	2	3	2	3	2	2	2	2	3	2	2	2	2	2	2	2	2	0.2
F4	2	2	2	2	2	2	2	2	2	4	2	2	5	3	5	2	3	1
	2	2	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.1
C4	2	2	2	2	2	2	2	2	2	4	2	5	4	2	5	3	3	1
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
P4	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
O2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
	2	2	2	2	2	2	2	2	3	2	2	2	2	2	2	2	2	0.1
F8	2	2	2	2	2	4	3	3	4	4	2	2	4	2	5	3	3	1
	2	2	2	2	2	3	2	2	2	2	2	2	2	2	2	2	2	0.1
T4	2	2	3	2	2	4	4	4	4	4	4	4	3	2	4	4	3	0.9
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
T6	2	2	2	2	2	2	2	2	3	2	2	2	3	2	2	2	2	0.1
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0
AV_j	2	2	2	2	2	3	3	3	3	3	3	3	3	2	3	3		
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2		
Var_j	0.1	0	0.2	0.9	0.5	0.8	0.6	0.6	0.8	0.9	0.7	1	0.9	0.2	1	0.5		
	0	0.1	0.1	0.1	0	0.1	0	0.1	0.1	0	0	0.1	0	0	0	0.1		

Note: SE is short for scalp electrode. For each cell of the table, the upper number corresponds to epileptic EEG time series and the lower number to normal EEG time series. AV_i stands for the average value of the *i* row embedding dimensions, and AV_j stands for the average value of the *j* line embedding dimensions. Var_i stands for the variance of the *i* row embedding dimensions, and Var_j stands for the variance of the *j* line embedding dimensions.

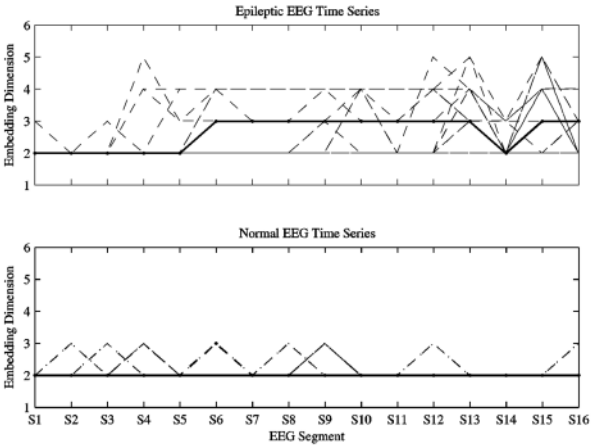


Fig. 5. Scheme of variation over time of EEG's embedding dimensions obtained by the differential entropy method. (Note: The curves marked by black points are constructed by AV_j values shown in Table 2.)

entropy method are $RMSE_{normal}$ = 0.0132 and $RMSE_{epileptic}$ = 0.0562. The results of variances and RMSE both indicate that the embedding dimensions of EEG signals from an epileptic vary over time more intensely during seizure than those from a normal person do. The average values of the embedding dimensions of normal and epileptic EEG signals are $ED_{normal} = 2$ and $ED_{epileptic} = 3$.

DISCUSSIONS

The results of the methods both show that the embedding dimension values of epileptic EEG signals vary intensively during seizure, whereas those of normal EEG signals basically remain stable; the embedding dimensions of EEG signals become much larger during seizure than those of normal EEG signals. The average value of the embedding dimensions of epileptic EEG signals is more than 2 times larger than those of normal EEG signals. Based on the results of Cao's method, the results of the differential entropy method also show that the embedding dimensions of epileptic EEG signals are larger than those of normal EEG signals. Although there is a great difference between the results of the two methods, namely, that the embedding dimension results obtained by the differential entropy method are much less than those obtained by Cao's method, the phenomena shown by the results of the two methods are consistent. The embedding dimensions of EEG signals during seizure become larger than those of normal EEG signals, and they vary briskly over time during seizure and have a bigger variation interval than those of normal EEG signals, whereas the embedding dimensions of normal EEG signals basically maintain stability.

An interesting phenomenon can be found in Fig. 2: normal EEG signals are random, whereas epileptic EEG signals are deterministic. For random data, based on Cao's method, the future values are independent of the past values; thus $E_2(d)$ will be equal to 1 for any d. For deterministic data, however, $E_2(d)$ is related to d and as a result cannot be a constant for all d; in other words, there must exist some d's such that $E_2(d) \neq 1$ [15]. For the left-hand plots corresponding to normal EEG signals in Fig. 2, except for 2 or 3 plots in which $E_2(d)$ varies slightly around 1, the others all show that after some slight variation at the beginning, $E_2(d)$ converges to 1 quickly. But we still cannot say that normal EEG signals are totally random, since $E_2(d)$ is not equal to 1 for all d's. For the right-hand plots corresponding to epileptic EEG signals in Fig. 2, we can say that epileptic EEG signals are strongly deterministic, since $E_2(d)$ in the right-hand plots remains unequal to 1 and varies briskly. The determinism of the epileptic EEG time series implies that some connection exists between future and past values of the epileptic EEG time series, and the EEG time series during seizure can be predicted. At present, arguments still exist about whether the EEG is random or deterministic, but the mainstream holds that the EEG is generated by a deterministic chaotic process [18]. The results of this paper verify that epileptic EEG time series are deterministic. Although these results show that the normal EEG time series are random, we cannot say that they are completely random, since different methods have different suitabilities for different conditions, such as low- or high-dimensional determinism [18], and $E_2(d)$ corresponding to normal EEG time series is not always equal to 1 (see Fig. 2). Based on the results of this paper, we can only determine that epileptic EEG time series are of stronger determinism than normal EEG time series. Moreover, a potential means for detecting epileptic seizure can also be found; that is, the difference of determinism between normal and epileptic EEG time series could be helpful for the detection of epileptic seizure.

If the human brain is considered as an NDS, the minimal embedding dimension gives the upper number of NDS freedom degrees and the minimal number of differential equations demanded for a mathematical modeling

of NDS. The results of embedding dimensions show that epileptic EEG time series are of a higher degree of freedom than normal EEG time series. Although we cannot obtain the accurate mathematical model of human brain NDS, there must be a relationship between the NDS and its output, and the property of an NDS can be showed by its output. Normal and epileptic EEG time series have different embedding dimensions, which also indicates that normal and epileptic may correspond to different NDS's, i.e., the NDS of human brain changes during seizure, which is meaningful for a mathematical modeling for EEG time series.

The results of the two methods used in this paper are quite different; we didn't discuss which method is more accurate because our purpose is to study the difference of embedding dimensions between normal and epileptic EEG signals. However, it is recommended to set the embedding dimensions of epileptic EEG signals at 17 and the embedding dimensions of normal EEG signals at 8 for the general processing of EEG time series. We are apt to believe that the results obtained by Cao's method are closer to the true values, since the embedding dimensions represent the minimal number of differential equations demanded for a mathematical modeling of NDS. The embedding dimension results of the differential entropy method are mainly around 2 or 3, which seem too small for a mathematical modeling of a real-world signal.

For the EEG data we used, not all the embedding dimension values of epileptic EEG segments were larger than those of normal EEG segments. We obtained 256 pairs of embedding dimension values of normal and epileptic EEG segments for the two methods. We then compared each pair of embedding dimension values $E_{ij}(i, j = 1,2,...,16)$ (the definition of E_{ij} can be found in *Results of Cao's method*) that are of the same indices, *i* and *j*. We found that for the results of Cao's method, there are 237 pairs $ED_{epileptic} > ED_{normal}$, 9 pairs $ED_{epileptic} = ED_{normal}$, and 10 pairs $ED_{epileptic} < ED_{normal}$. For the results of the differential entropy method, in a similar way we found that the statistical results are 96, 152, and 8, respectively. (The embedding dimension values in detail obtained by the two methods are shown in Tables 1 and 2.) The phenomenon $ED_{epileptic} < ED_{normal}$ exist for both methods; however, it is believed that it cannot prevent us from getting the impression that the embedding dimension of EEG time series during seizure becomes larger than that of the normal EEG time series, since only 3.91% and 3.13% out of 256 pairs of embedding dimension values show $ED_{epileptic} < ED_{normal}$ for Cao's method and differential entropy method, respectively.

Based on the results of this paper, normal and epileptic EEG time series show different embedding dimensions and degree of determinism; therefore, we consider that embedding dimensions, and even degree of determinism, have potential application value in the clinical diagnosis of epilepsy through further research. In future work, we will further investigate the differences of embedding dimensions among EEG data from different types of seizures and different monitored subjects (including different ages, sexes, and other variables).

CONCLUSIONS

Cao's method and the differential entropy method were applied to compute the minimum embedding dimensions of normal and epileptic EEG time series, respectively. The phenomena shown by the results of these two methods are consistent: the embedding dimensions of epileptic EEG signals is much larger than those of normal EEG signals, even up to 2 times, according to the results of Cao's method. The embedding dimensions of EEG signals varies briskly during seizure, whereas the embedding dimensions of normal EEG signals basically maintain stability. The embedding dimensions of epileptic EEG signals have a bigger variation interval than those of normal EEG signals. The variation of the embedding dimensions of EEG signals during seizure also means that the variation of the freedom degree of the human brain NDS increases during seizure. Furthermore, based on the results of Cao's method, we found that normal EEG signals are to some degree random, whereas epileptic EEG signals are strongly deterministic, which implies that epileptic EEG signals can be predicted well. Therefore it is proposed that the embedding dimension can be a supplementary parameter for the characterization of epileptic seizures.

APPENDIX

1. Cao's method

Supposing that $x_1, x_2, \cdots, x_N$ is a time series, the embedding dimension can be determined by Cao's method as follows:

Reconstruct the time series as time delay vectors in phase space:

$$Y_i(d) = [x(i), x(i+\tau), x(i+2\tau), \cdots, x(i+(d-1)\tau)], \quad i = 1,2,\cdots,N-(d-1)\tau \tag{A.1}$$

where d is the embedding dimension and τ is the time delay. $Y_i(d)$ is the i reconstructed vector and $Y_i^{NN}(d)$ is the nearest neighbor of $Y_i(d)$ in embedding dimension d as follows:

$$Y_i^{NN}(d) = [x^{NN}(i), x^{NN}(i+\tau), x^{NN}(i+2\tau), \cdots, x^{NN}(i+(d-1)\tau)] \tag{A.2}$$

Define

$$a_2(i,d) = \frac{\left\| Y_i(d+1) - Y_i^{NN}(d+1) \right\|}{\left\| Y_i(d) - Y_i^{NN}(d) \right\|} \tag{A.3}$$

where $\|\cdot\|$ is the Euclidian distance and is given by the maximum norm here. $Y_i(d)$ is the i reconstructed vector, and $Y_i^{NN}(d+1)$ is its nearest neighbor in embedding dimension $d+1$. The mean value of all $a_2(i, d)$'s is defined as

$$E(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} a_2(i,d) \quad \text{(A.4)}$$

$E(d)$ is dependent only on the dimension d and lag τ. To investigate its variation from d to $d + 1$, define

$$E_1(d) = E(d+1) / E(d) \quad \text{(A.5)}$$

$E_1(d)$ will stop changing when d is greater than some value d_0, which is the minimum embedding dimension we look for. In *Results of Cao's method*, we take the first plot in the left side in Fig. 2 as an example for explaining how to determine the embedding dimension value of a time series by Cao's method; at this time the embedding dimension value of the EEG signal segment is $d_0 = 10$.

Another quantity is defined to distinguish deterministic signals from stochastic signals. Let

$$E^*(d) = \frac{1}{N - d\tau} \sum_{i=1}^{N-d\tau} \left| x(i + d\tau) - x^{NN}(i + d\tau) \right| \quad \text{(A.6)}$$

$$E_2(d) = E^*(d+1) / E^*(d) \quad \text{(A.7)}$$

$E_1(d)$ is calculated for determining the minimum embedding dimension of time series, and $E_2(d)$ for distinguishing deterministic data from random data.

2. Differential entropy method

The differential entropy method can be summarized as below:

(a) For the given signal $x(t)(t = 1,2,...,N)$, generate its N_s surrogates $x_{s,i}(t), i = 1,...,N$ by performing a random permutation of the time samples.

(b) The Kozachenko-Leonenko estimates for the time delay embedded versions of the original time series $H(x,m,\tau)$ and its surrogates $H(x_{s,i},m,\tau)$ are computed using Eq. 8 for increasing m and τ.

$$H(x) = \sum_{j=1}^{N} 1\mathrm{n}(N\rho_j) + 1\mathrm{n}\ 2 + C_E \quad \text{(A.8)}$$

(c) Minimize the ratio

$$I(m,\tau) = \frac{H(x,m,\tau)}{\left\langle H(x_{s,i},m,\tau) \right\rangle_i} \quad \text{(A.9)}$$

to determine the optimal embedding parameters, where $\langle \cdot \rangle_i$ denotes the average over i.

(d) To penalize for higher embedding dimensions, the minimum description length (MDL) method is superimposed, yielding the entropy ratio (ER):

$$R_{ent}(m,\tau) = I(m,\tau)(1 + \frac{m\ 1\mathrm{n}\ N}{N}) \quad \text{(A.10)}$$

where N is the number of delay vectors.

(e) The minimum of the plot of the entropy ratio yields the optimal set of embedding parameters, namely, optimal embedding dimension m and delay time τ.

The authors thank Ralph G. Andrzejak for many helpful discussions. This research is supported by the National Nature Science Foundation of China under Grant Nos. 40574051 and 40774054.

REFERENCES

1. McGrogan N. Neural network detection of epileptic seizures in the electroencephalogram. Oxford University; 1999. Available online: http://citeseer.ist.psu.edu/625725.html.
2. Iasemidis LD, Principe JC, Sackellares JC. Measurement and quantification of spatio-temporal dynamics of human epileptic seizures. In: Akay M, editor. Nonlinear signal processing in medicine. IEEE Press; 1999. p. 1-27.
3. Niedermeyer E, Lopes da Silva F. Electroencephalography: Basic principles, clinical applications, and related fields. Baltimore: Lippincott Williams & Wilkins; 1993.
4. Bauer M, Heng H, Martienssen W. Characterization of spatiotemporal chaos from time series. Phys Rev Lett. 1993;71:521-4.
5. Richman JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol. 2000;278:H2039-49.
6. Gautama T, Mandic DP, Van Hulle MM. Indications of nonlinear structures in brain electrical activity. Phys Rev E Stat Nonlin Soft Matter Phys. 2003;67:046204.
7. Saito K, Suzuki H, Kawakami Y. Power spectrum density of EEGs of sleeping epilepsy-prone El mice and their non-epileptic mother strain. J Physiol Sci. 2006;56:313-6.
8. Phothisonothai M, Nakagawa M. Fractal-based EEG data analysis of body parts movement imagery tasks. J Physiol Sci. 2007;57:217-26.
9. Li X, Ouyang G. Nonlinear similarity analysis for epileptic seizures prediction. Nonlinear Analysis. 2006;64:1666-78.
10. Diambra L, Malta CP. Nonlinear models for detecting epileptic spikes. Phys Rev E. 1999;59:929-37.
11. Silva C, Pimentel IR, Andrade A, Foreid JP, Ducla-Soares E. Correlation dimension maps of EEG from epileptic absences. Brain Topogr. 1999;11:201-9.
12. Kannathal N, Puthusserypady SK, Choo Min L. Complex dynamics of epileptic EEG. Conf Proc IEEE Eng Med Biol Soc. 2004;1:604-7.
13. Frank GW, Lookman T, Nerenberg MAH, Essex C, Lemieux J, Blume W. Chaotic time series analyses of epileptic seizures. Physica D. 1990;46:427-38.
14. 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. Nat Med. 1998;4:1173-6.
15. Cao LY. Practical method for determining the minimum embedding dimension of a scalar time series. Physica D. 1997;110:43-50.
16. Gautama T, Mandic DP, Van Hulle MM. A differential entropy based method for determining the optimal embedding parameters of a signal. In: Proceedings of the Int Conf on Acoustics, Speech and Signal Processing [C], Hong Kong. 2003;35:29-32.
17. Fraser AM, Swinney HL. Independent coordinates for strange attractors from mutual information. Phys Rev A. 1986;33:1134-40.
18. Jeong J, Kim MS, Kim SY. Test for low-dimensional determinism in electroencephalograms. Phys Rev E Stat Phys Plasmas Fluids Relat Interdiscip Topics. 1999;60:831-7.