

# ELECTROENCEPHALOGRAM ANALYSIS WITH APPROXIMATE ENTROPY TO HELP IN THE DIAGNOSIS OF ALZHEIMER'S DISEASE

Daniel Abásolo<sup>1</sup>, Roberto Hornero<sup>1</sup>, Pedro Espino<sup>2</sup>, Alonso Alonso<sup>1</sup>, Ramón de la Rosa<sup>1</sup>

<sup>1</sup>E.T.S. Ingenieros de Telecomunicación, University of Valladolid

Paseo del Cementerio s/nº, 47011 - Valladolid (Spain)

Phone: +34 983 423000 ext. 25569 Fax: +34 983 423667 E-mail: danaba@tel.uva.es

<sup>2</sup>Hospital Universitario de Valladolid

Avda. Ramón y Cajal 3, 47005 - Valladolid (Spain)

**Abstract** - Alzheimer's disease (AD) is the main cause of dementia in western countries. Although a definite diagnosis of this illness is only possible by necropsy, the analysis of non-linear dynamics in electroencephalogram (EEG) signals could help physicians in this difficult task. In this study we have applied Approximate Entropy (ApEn) to analyze the EEG background activity of patients with a clinical diagnosis of Alzheimer's disease and control subjects. ApEn is a newly introduced statistic that can be used to quantify the complexity (or irregularity) of a time series. We have divided the EEG data into frames to calculate their ApEn. Our results show that the degree of complexity of EEGs from control subjects is higher. Applying the ANOVA test, we have verified that there was a significant difference ( $p < 0.05$ ) between the EEGs of these groups.

**Keywords** - Alzheimer's Disease (AD), Electroencephalogram (EEG), Approximate entropy (ApEn), Complexity

## I. INTRODUCTION

Alzheimer's disease (AD) is a degenerative neurological disease considered as the main cause of dementia in western countries [1]. Clinically, AD manifests as a slowly progressive impairment of mental functions whose course lasts several years prior to death. The brain of AD patients shows a diffuse atrophy of the cortex and, microscopically, there exist neuritic plaques containing amyloid  $A\beta$ , neurofibrillary tangles and deposits of amyloid in the walls of the brain arteries. Although a definite diagnosis is only possible by necropsy, a differential diagnosis with other types of dementia and with major depression should be attempted. Magnetic resonance imaging and computerized tomography can be normal in the early stages of AD but a diffuse cortical atrophy is the main sign in brain scans. Mental status tests are also useful.

In these patients, the electroencephalogram (EEG) shows generalized changes with a diffuse slowing of the background activity [2]. Nevertheless, in the early stages of the disease, the EEG may exhibit normal frequencies. In several studies, the use of quantitative studies of the EEG did not provide an important advantage – compared to the visual analysis – when trying to differentiate AD patients from controls and from those suffering a major depression [3]. However, the analysis of non-linear dynamics in EEG signals could improve medical diagnosis as well as the understanding of physiological processes and their changes due to disease. With this approach, new techniques have

been applied to quantify differences in the EEG series. For example, several stages of the human EEG were defined according to the value of their correlation dimension ( $CD$ ). The analyses were performed on EEGs recorded from normal individuals during deep sleep [4] and wakefulness with alpha-wave activity [5]. It was shown that the correlation dimension of the chaotic attractors decreases as the brain activity switches from alpha waves ( $CD = 6.1 \pm 0.5$ ) to deep sleep ( $CD = 4.4 \pm 0.1$ ). In other studies, the correlation dimension was estimated from the EEGs of patients suffering from petit mal epilepsy [6] and the Creutzfeldt-Jakob disease [7]. The results showed a decrease in the correlation dimension of the two pathological states when compared with the above results ( $CD = 3.8 \pm 0.2$  and  $CD = 2.05 \pm 0.09$ , respectively from Creutzfeldt-Jakob encephalopathy and petit mal epilepsy). Thus, the correlation dimension analysis furnishes an elegant method for assessing and classifying various dynamic states of the brain. In previous works, we have already applied this non-linear method to assess cognitive function in schizophrenia [8], and to study the EEG of epileptic patients [9].

However, there are some drawbacks in this non-linear method that should be considered. The most important one is that the Grassberger and Procaccia algorithm used to estimate the  $CD$  assumes the time series to be stationary [10], and this is generally not true with biological data.

In the study reported here, we have applied a different method – approximate entropy (ApEn) – without the shortcomings of the aforementioned one. The ApEn has been used to quantify the regularity in the EEG background activity of 7 patients with a clinical diagnosis of AD and 7 control subjects and to determine whether there are differences or not between both groups.

## II. METHODOLOGY

We have studied 7 patients (1 man and 6 women) with a diagnosis of AD based upon history, clinical examination, brain scans and a Mini-Mental State Examination (MMSE) [11], generally accepted as a quick and simple way to evaluate cognitive function. Patients with a mean age of  $75.6 \pm 6.0$  were recruited from the University Hospital of Valladolid, Spain. The control group of subjects without past or present neurological disorders included 7 people (2 men and 5 women) with a mean age of  $58.7 \pm 3.5$ .

Electrodes were placed in accordance with the 10-20 system for the EEG recording. All electrodes were referenced to the chin. The subjects were awake, relaxed, in

---

This work has been supported by the grant project of Junta de Castilla y León VA090/02.

a quiet state with closed eyes. EEG data were first processed with a low-pass hardware filter at 100 Hz. Then, they were sampled at 256 Hz, digitized to 12 bits, and digitally filtered with a high-pass filter at 0.4 Hz and a low-pass filter at 70 Hz. A 50 Hz notch filter was also used to eliminate the noise owing to the electrical mains. We recorded more than five minutes of EEG from each subject. Firstly, a specialist physician discarded the noisy parts at the beginning of the recording and those corresponding to movements. This analysis was done visually on all the EEG channels. Besides, EEGs were organized in frames of 5 seconds (1280 points). Those frames began when the recording was stable (i.e. the noisy parts at the beginning of the recording were discarded).

In the analysis of the EEG background activity, we could analyze EEG data corresponding to any lead. However, we used the temporal series obtained from the parietal electrode P3 following the advice of an electroencephalographer, because there are less artifacts and the rhythmical activity is more patent.

The correlation dimension (*CD*) is a method that has been used to characterize the regularity of the systems and time series. It is a measure of the complexity of the process being investigated, which characterizes the distribution of points in the phase space. The larger the *CD* of the attractor, the more complicated the behavior of the non-linear system. The traditional Grassberger and Procaccia algorithm [10] assumes the time series to be stationary and noise free. However, these assumptions are not generally true. The stationarity constraints imposed on the time series data results in the reduction of the scaling region, making the calculation of *CD* error prone [12]. Hence, a different algorithm is necessary to study with more accuracy the EEG background activity.

Approximate Entropy (ApEn) is a recently formulated family of parameters and statistics quantifying regularity (orderliness) in serial data. It was first proposed by Pincus in 1991 [13] and has been used mainly in the analysis of heart rate variability [14-15], endocrine hormone release pulsatility [16], estimating regularity in epileptic seizure time series data [12] and estimating the depth of anesthesia [17]. ApEn is scale invariant and model independent, evaluates both dominant and subdominant patterns in data, and discriminates series for which clear feature recognition is difficult. Notably it detects changes in underlying episodic behavior not reflected in peak occurrences or amplitudes [18].

ApEn was introduced as a quantification of regularity in sequences and time series data, initially motivated by applications to relatively short, noisy data sets [13]. It assigns a non-negative number to a sequence or time series, with larger values corresponding to more instances of recognizable features or patterns in the data. When applied to EEG time series, larger values indicate higher complexity [19]. Two input parameters, a run length  $m$  and a tolerance window  $r$ , must be specified to compute ApEn. Briefly, ApEn measures the logarithmic likelihood that runs of patterns that are close (within  $r$ ) for  $m$  contiguous

observations remain close (within the same tolerance width  $r$ ) on subsequent incremental comparisons. It is imperative to consider  $\text{ApEn}(m,r)$  – or  $\text{ApEn}(m,r,N)$ , where  $N$  is the number of points of the time series – as a *family* of parameters: comparisons are intended with fixed  $m$  and  $r$  [20].

Formally, given  $N$  data points from a time series  $\langle x(n) \rangle = x(1), x(2), \dots, x(N)$ , two input parameters,  $m$  and  $r$ , must be fixed to compute ApEn, denoted precisely by  $\text{ApEn}(m,r,N)$ . To define ApEn, one should follow these steps:

- 1) Form  $m$ -vectors  $X(1) \sim X(N-m+1)$  defined by:  $X(i) = [x(i), x(i+1), \dots, x(i+m-1)]$   $i = 1 \sim N-m+1$ . (1) These vectors represent  $m$  consecutive  $x$  values, commencing with the  $i$ th point.
- 2) Define the distance between  $X(i)$  and  $X(j)$ ,  $d[X(i), X(j)]$ , as the maximum absolute difference between their respective scalar components, i.e.,

$$d[X(i), X(j)] = \max_{k=1,2,\dots,m} (|x(i+k-1) - x(j+k-1)|) \quad (2)$$

- 3) For a given  $X(i)$ , count the number of  $j$  ( $j = 1 \sim N-m+1$ ,  $j \neq i$ ) such that  $d[X(i), X(j)] \leq r$ , denoted as  $N^m(i)$ . Then, for  $i = 1 \sim N-m+1$ ,

$$C_r^m(i) = N^m(i) / (N-m+1) \quad (3)$$

The  $C_r^m(i)$  values measure *within a tolerance  $r$*  the regularity, or frequency, of patterns similar to a given one of *window length  $m$* .

- 4) Compute the natural algorithm of each  $C_r^m(i)$ , and average it over  $i$ ,

$$\phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_r^m(i) \quad (4)$$

- 5) Increase the dimension to  $m+1$ . Repeat steps 1) to 4) and find  $C_r^{m+1}(i)$  and  $\phi^{m+1}(r)$ .
- 6) We define approximate entropy by:

$$\text{ApEn}(m,r,N) = \phi^m(r) - \phi^{m+1}(r) \quad (5)$$

As suggested by Pincus, for the study discussed in this paper ApEn was calculated with the widely established parameter values of  $m=2$  and  $r$  a fixed value between 0.1 and 0.25 of the standard deviation (SD) of the original data sequence  $\langle x(n) \rangle$ . *Normalizing  $r$*  to each time series SD in this manner gives ApEn a translation and scale invariance, in that it remains unchanged under uniform process magnification, reduction, or constant shift to higher or lower values [20]. Several studies [13,18,21] have demonstrated that these input parameters produce good statistical reproducibility for ApEn for time series of length  $n \geq 60$ , as considered herein.

Because of the nonlinear character of EEG signals, ApEn can be used as a powerful tool in the study of the EEG background activity.

We have calculated ApEn using a short computer program, developed with the MATLAB® program.

### III. RESULTS

As we have mentioned before, prior to the estimation of the ApEn from EEG data of patients with AD and control subjects, the data were reviewed by a specialist physician and artifact-free segments from the P3 channel were chosen. The data were then divided in frames of 1280 points (5 seconds from the EEG recording sampled at 256 Hz).

ApEn( $m, r, N$ ) has to be seen as a family of parameters, where  $N$  is the number of points (1280). As suggested by Pincus in [20], we have used  $m=2$  and four different values of  $r$ : 0.1, 0.15, 0.2 and 0.25 of the standard deviation of the data. The results of the approximate entropy have been averaged based on all the artifact-free frames within the five-minute period of EEG recordings of control subjects and patients with a diagnosis of AD.

In Table 1 we show the mean ApEn(2,0.25,1280) values calculated from the EEG frames evaluated for each subject. The results were higher for the control subjects' EEGs, with a mean value of  $0.93 \pm 0.11$ , compared to those obtained from patients' EEGs, where the mean value was  $0.78 \pm 0.10$ . The mean values of the approximate entropy were also higher for control subjects with  $r = 0.1$ ,  $r = 0.15$  and  $r = 0.2$  of the standard deviation of the data.

These results indicate that the degree of complexity of EEGs corresponding to the control subjects is higher (higher ApEn values) than the one of EEGs from patients with AD. Moreover, we have applied the ANOVA test [22], where we have used the mean ApEn values calculated from the EEG frames of 7 control subject and 7 patient with AD. We have verified with this analysis that there was a significant difference ( $F_{1,12} = 6.89$ ;  $p = 0.022$ ) between the EEGs of these groups. These differences between both groups were also evident with the other values of  $r$ .

**TABLE I**

ApEn(2,0.25,1280) mean values of the EEG frames from control subjects and patients with Alzheimer's disease.

PATIENTS WITH ALZHEIMER'S DISEASE			CONTROL SUBJECTS		
<i>Id.</i>	<i>Age-Sex</i>	<i>ApEn</i>	<i>Id.</i>	<i>Age-Sex</i>	<i>ApEn</i>
alz-1	80-F	0.8449	con-1	57-F	0.8842
alz-2	66-F	0.8230	con-2	59-F	0.8492
alz-3	83-M	0.9386	con-3	59-F	0.8429
alz-4	73-F	0.6684	con-4	64-F	0.8995
alz-5	76-F	0.6544	con-5	54-M	1.0993
alz-6	80-F	0.7906	con-6	56-F	0.8611
alz-7	71-F	0.7656	con-7	62-M	1.0575
MEAN VALUE	$75.6 \pm 6.0$	$0.78 \pm 0.10$	MEAN VALUE	$58.7 \pm 3.5$	$0.93 \pm 0.11$

### IV. DISCUSSION

We have analyzed the EEG background activity of 7 control subjects and 7 patients with AD by means of nonlinear methods. Several properties of ApEn facilitate its utility for empirical time series analysis of the sort of EEGs [20]:

- 1) ApEn is nearly unaffected by noise below a *de facto* specified filter level ( $r$ ).
- 2) ApEn can be applied to time series of 50 or more points with good reproducibility.
- 3) ApEn is finite for stochastic, noisy deterministic and composite processes, these last of which are likely models for complicated biological data sets.
- 4) Increasing values of ApEn correspond to intuitively increasing process complexity.

Our results show that the ApEn is higher in the control subjects' EEGs when compared with the ApEn values of the EEGs of patients with a diagnosis of AD. We have previously mentioned that in some studies, the use of quantitative studies of the EEG did not provide an important advantage compared to the visual analysis when trying to differentiate AD patients from control subjects [3]. However, and as the ApEn parameter can be viewed as a measure of the complexity of a system, these results suggest that the degree of complexity of the control subjects' EEGs is higher than the one corresponding to the AD patients. The ANOVA test shows that the differences between both groups are significant ( $p$ -ANOVA  $< 0.05$ ). This confirms findings associated with the fact that a diffuse slowing of the background activity may be found in the EEGs of patients with AD [2].

Some remarks must be made in the interpretation of these results. First of all, our conclusions are based on analyses of variance (ANOVA) comparing EEG signals of patients and control subjects, and not on the absolute values of the approximate entropy. Secondly, the differences in ApEn values are significant when patients are considered as a group. For the diagnosis of a single case, it would be necessary to have a limit value of normalcy. Even though there are differences in the middle edge of the two populations, all subjects were adults with an advanced age and we do believe that these differences do not bias the results.

On the other hand, the detected reduction of EEG complexity is not specific for AD. It appears in several physiological and pathological states, including sleep, schizophrenia and other neurobiological disorders. Thus, further work with more patients and control subjects must be carried out to establish the possible usefulness of this evaluation.

### V. CONCLUSION

Approximate entropy is a parameter that can give robust estimate from short data and is attractive for dynamic analysis. There are of course other nonlinear parameters that can be applied to short data, but it appears that ApEn has potential widespread utility to practical data analysis and clinical application due to the salient features it bears. Increasing ApEn corresponds to intuitively increasing process complexity.

Our experimental results prove the potential applications of ApEn in reflecting differences in the

complexity of EEG data time series of patients with a diagnosis of Alzheimer's disease and control subjects. The results indicate that the degree of complexity of EEGs corresponding to the control subjects is higher (higher ApEn values). However, this reduction of complexity is not specific to AD and further work is necessary to establish the possible usefulness of this evaluation in the medical diagnosis.

Possible future lines of research include the non-linear analysis of multiple EEG channels, the study of correlations with other psychopathological scales and other assessment instruments.

In spite of the above-mentioned precautions, we suggest that the non-linear analysis of EEG data might be useful to help physicians in the diagnosis of the Alzheimer's disease.

#### ACKNOWLEDGMENT

We are very grateful to María Luisa Baz, Ph.D., who kindly revised the English version of this paper.

#### REFERENCES

- [1] T.D. Bird, "Alzheimer's disease and other primary dementias", in *Harrison's Principles of Internal Medicine*, E. Braunwald, A.S. Fauci, D.L. Kasper, S.L. Hauser, D.L. Longo and J.L. Jameson (Eds.), New York: The McGraw-Hill Companies, Inc., 2001, pp. 2391-2399.
- [2] O.N. Markand, "Organic brain syndromes and dementias", in *Current Practice of Clinical Electroencephalography*, D.D. Daly, T.A. Pedley (Eds.), New York: Raven Press, Ltd, 1990, pp. 401-423.
- [3] R.P. Brenner, C.F. Reynolds 3<sup>rd</sup> and R.F. Ulrich, "Diagnostic efficacy of computerized spectral versus visual EEG analysis in elderly normal, demented and depressed subjects", *Electroencephalogr. Clin. Neurophysiol.*, vol. 69, pp. 110-117, 1988.
- [4] A. Babloyantz and J.M. Salazar, "Evidence of chaotic dynamics of brain activity during the sleep cycle", *Phys. Lett. A*, vol. 111, pp. 152-156, 1985.
- [5] A. Babloyantz and A. Destexhe, "Strange attractors in the human cortex", in *Temporal Disorders in Human Oscillatory Systems*, L. Rensing, U. van der Heiden and M.C. Mackey (Eds.), *Springer Series in Synergetics*, vol. 36, Berlin, Heidelberg, New York: Springer Verlag, 1987, pp. 48-56.
- [6] A. Babloyantz and A. Destexhe, "Low-dimensional chaos in an instance of epilepsy", *Proc. Natl. Acad. Sci. USA*, vol. 83, pp. 3513-3517, 1986.
- [7] A. Babloyantz and A. Destexhe, "The Creutzfeldt-Jakob disease in the hierarchy of chaotic attractors", in *From chemical to biological organisation*, M. Markus, S. Müller and G. Nicolis (Eds.), *Springer Series in Synergetics*, vol. 39, Berlin, Heidelberg, New York: Springer Verlag 1988, pp. 307-316.
- [8] R. Hornero, A. Alonso, N. Jimeno, A. Jimeno and M. López, "Estimation of correlation dimension to evaluate cognitive performance in schizophrenic patients using a new computer technique", *Nonlinear Dynamics, Psychology and Life Sciences*, vol. 3, pp 49-63, 1999.
- [9] R. Hornero, P. Espino, A. Alonso and M. López, "Estimating complexity from EEG background activity of epileptic patients", *IEEE Eng. Med. Biol.*, vol. 18, pp. 73-79, 1999.
- [10] P. Grassberger and I. Procaccia, "Characterization of strange attractors", *Phys. Rev. Lett.*, vol. 50, pp. 346-349, 1983.
- [11] M.F. Folstein, S.E. Folstein and P.R. McHugh, "Minimal state. A practical method for grading the cognitive state of patients for the clinician", *J. Psychiatr. Res.*, vol. 12, pp. 189-198, 1975.
- [12] N. Radhakrishnan, B.N. Gangadhar, "Estimating regularity in epileptic seizure time-series data. A complexity-measure approach", *IEEE Eng. Med. Biol.*, vol. 17 (3), pp. 89-94, 1998.
- [13] S.M. Pincus, "Approximate entropy as a measure of system complexity", *Proc. Natl. Acad. Sci. USA*, vol. 88, pp. 2297-2301, 1991.
- [14] S.M. Pincus, "Heart rate control in normal and aborted-SIDS infants", *Am. J. Physiol. (Regulatory Integrative Comp. Physiol.)*, pp. R638-R646, 1991.
- [15] S.M. Pincus, "Physiological time series analysis: what does regularity quantify?", *Am. J. Physiol.*, vol. 266 (*Heart Circ. Physiol.*), pp. H1643-H1656, 1994.
- [16] S.M. Pincus, "Older males secrete luteinizing hormone and testosterone more irregularly and joint more asynchronously, than younger males", *Proc. Natl. Acad. Sci. USA*, vol. 93, pp. 14100-14105, 1996.
- [17] X.-S. Zhang and R.J. Roy, "Derived fuzzy knowledge model for estimating the depth of anesthesia", *IEEE Trans. Biomed. Eng.*, vol. 48 (3), March 2001.
- [18] S.M. Pincus and D.L. Keefe, "Quantification of hormone pulsatility via an approximate entropy algorithm", *Am. J. Physiol.*, vol. 262, pp. E741-E754, 1992.
- [19] H. Bo, Y. Fusheng, T. Qingyu and C. Tin-cheung, "Approximate entropy and its preliminary application in the field of EEG and cognition", in *Proc. Of the 20<sup>th</sup> Annual International Conference of the IEEE Eng. Med. Biol. Society*, vol. 20 (4), 1998, pp. 2091-2094.
- [20] S.M. Pincus, "Assessing serial irregularity and its implications for health", *Ann. NY Acad. Sci.*, vol. 954, pp. 245-267, 2001.
- [21] D.T. Kaplan, M.I. Furman, S.M. Pincus, S.M. Ryan, L.A. Lipsitz and A.L. Goldberger, "Aging and the complexity of cardiovascular dynamics", *Byophys. J.*, vol. 59, pp. 945-949, 1991.
- [22] A.O. Allen, *Probability, statistics, and queuing theory with computer science applications*, 2nd ed. San Diego, Academic Press, 1990.