Evaluation of Resting-State Magnetoencephalogram Complexity in Alzheimer's Disease with Multivariate Multiscale Permutation and Sample Entropies

Evaluation of Resting-State Magnetoencephalogram Complexity in Alzheimer's Disease with Multivariate Multiscale Permutation and Sample Entropies

Hamed Azami[1], *Student Member, IEEE*, Keith Smith[1,2], *Student Member, IEEE*, Alberto Fernandez[3], and Javier Escudero[1,*], *Member, IEEE*

***Abstract*— Alzheimer's disease (AD) is one of the fastest growing neurological diseases in the world. We evaluate multivariate multiscale sample entropy (mvMSE) and multivariate multiscale permutation entropy (mvMPE) approaches to distinguish resting-state magnetoencephalogram (MEG) signals of 36 AD patients from those of 26 normal controls. We also discuss about choosing the appropriate embedding dimension value as an effective parameter for mvMPE and MPE for the first time. The results illustrate that both the mvMPE and mvMSE can be useful in the diagnosis of AD, although with different running times and abilities. In addition, our findings show that the MEG complexity analysis performed on deeper time scales by mvMPE and mvMSE may be a useful tool to characterize AD. In most scale factors, the average of the mvMPE and mvMSE values of AD patients are lower than those of controls.**

I. Introduction

Alzheimer's disease (AD), which is the most common type of dementia, is a neurodegenerative disorder affecting more than 30 million people in the world [1]. Changes in biomedical signals, such as electroencephalogram (EEG) and magnetoencephalogram (MEG), have been investigated widely for the diagnosis of AD in the few past decades [2, 3].

Nonlinear methods, such as entropy-based techniques, fractal dimension, and Lempel–Ziv complexity of MEG and EEG recordings, showing the regularity or complexity of signals, are helpful and widely used to characterize the brain activity of AD patients. Research illustrates that EEG and MEG background activity is less complex and more regular in AD patients than in healthy control subjects [3, 4].

One of the most important approaches for quantifying the degree of regularity of signals is entropy [5]. Sample entropy (SaE) [6] and permutation entropy (PE) [7] are two popular entropy metrics. Both PE and SaE have their own advantages and disadvantages [2]. PE is considerably faster than SaE although SaE is more flexible [7]. However, these entropies are limited to evaluating the values of entropy for only one temporal scale, the one associated with the original sampling of the signals. This may limit the ability of these to inspect dynamics residing at longer temporal scales.

In this sense, multiscale (sample) entropy (MSE) [8] and multiscale PE (MPE) [9] were proposed to calculate entropy over a range of scales to evaluate the complexity of a time series. In fact, multiscale-based approaches offer much more information. Since biomedical signals are often acquired by many channels, multivariate MSE (mvMSE) [5] and multivariate MPE (mvMPE) [9] have been recently proposed. In [9] and [2], the effects of AD on the mvMPE and mvMSE, respectively, of EEG signals were studied. Although both EEG and MEG recordings have high temporal resolution, MEG signals have some advantages over the EEG signals. The MEG recordings are not related to any reference point and they are less influenced by extra-cerebral tissues than the EEGs [10]. In addition, the number of subjects was low [2, 9]. This may affect the reliability of the results. Moreover, in those papers [2, 9], only frontal and occipital areas, over eight channels, were considered and the role of the appropriate embedding dimension, a vital parameter in PE-based methods, was not discussed.

To tackle these limitations, we investigate mvMPE and mvMSE of 62 subjects' MEG signals (36 AD patients and 26 controls) for five main regions, including anterior, central, right and left lateral, and posterior areas, using eight channels for each region.

II. Materials

A. Subject Groups

All 62 subjects gave their informed agreement for the study, which was approved by the local ethics committee. Diagnoses were confirmed with thorough tests. To screen the cognitive status, the mini-mental state examination (MMSE) was utilized [11].

The 36 AD patients (24 women; age = 74.06 ± 6.95 years, mean± standard deviation, SD; MMSE score = 18.06 ± 3.36, mean±SD) met the criteria for probable AD according to the guidelines of the NINCDS-ADRDA [12].

The CON group was formed by 26 subjects (17 women; age = 71.77 ± 6.38 years; MMSE score = 28.88 ± 1.18, mean±SD). The difference in age between two groups was not significant (p-value = 0.1911, Student's t-test).

B. MEG Data

Resting state MEG data were obtained with a 148-channel whole-head magnetometer (MAGNES 2500 WH, 4D Neuroimaging) in a magnetically shielded room at the MEG Centre Dr. Pérez-Modrego (Spain). The subjects laid on a

[1] Hamed Azami, Keith Smith and Javier Escudero* are with the Institute for Digital Communications, School of Engineering, The University of Edinburgh, Edinburgh, King's Buildings, EH9 3JL, United Kingdom. (Phone: +44 131 650 5599, emails: hamed.azami@ed.ac.uk, k.smith@ed.ac.uk, javier.escudero@ed.ac.uk). * *Corresponding Author.*

[2] Keith Smith is also with the Alzheimer Scotland Dementia Research Centre, Psychology and Language Sciences, University of Edinburgh, 7 George Square, Edinburgh, EH8 9JZ.

[3] Alberto Fernandez is with the Departamento de Psiquiatría Psicología Médica, Universidad Complutense de Madrid, Madrid, Spain. He is also with Laboratorio de Neurociencia Cognitiva y Computacional, Centro de Tecnología Biomédica, Universidad Politécnica de Madrid and Universidad Complutense de Madrid, Madrid, Spain and with the Insitituto de Investigación Sanitaria San Carlos (IdSSC).

hospital bed in a relaxed state with eyes closed. They were asked to avoid falling asleep and not to move head and eyes. For each participant, five minutes of MEG resting state activity were recorded at a sampling frequency (f_s) of 169.54Hz. The signals were divided into segments of 10s (1695 samples per channel) and visually inspected using an automated thresholding procedure to discard segments significantly contaminated with artefacts [11]. The effect of cardiac artifact was reduced from the recordings using a constraint blind source separation procedure [13] to avoid bias in the computation of multivariate techniques. Finally, a bandpass FIR filter with cut-offs at 1.5Hz and 40Hz was applied to the data.

III. Methods

A. Multiscale Entropy

Both MSE and MPE include two main steps:

1. First, a "coarse-graining" process is applied to a P-variate (channel) signal $\mathbf{Y} = \{y_{k,b}\}_{b=1}^{C}$, k=1,…,P and C is the length of the each signal of each channel. According to (1), each element of the coarse-grained time series is defined as:

$$x_{k,i}^{(\beta)} = \frac{1}{\beta}\sum_{b=(i-1)\beta+1}^{i\beta} y_{k,b} \qquad 1 \le i \le \left\lfloor \frac{C}{\beta} \right\rfloor = N \tag{1}$$

where β is the time scale factor.

2. In the second step, for a defined scale factor, the mvSE or mvPE is calculated [5, 9].

B. Multivariate Entropy Methods

1) Multivariate Permutation Entropy

For each scale factor β and for each channel k=1,2,…,P, assume the coarse-grained time series is $\mathbf{z} = \{x_{k,1}, x_{k,2}, ..., x_{k,N}\}$. At each sample t of $\mathbf{z}$, a vector including the d-th subsequent values is constructed as: $\{x_{k,t}, x_{k,t+l}, ..., x_{k,t+(d-2)l}, x_{k,t+(d-1)l}\}$ for t=1,2,…,N-(d-1)l, where d, named the embedding dimension, determines how much information is contained in each vector and l is the time delay. To calculate the PE, the d values $\{x_{k,t}, x_{k,t+l}, ..., x_{k,t+(d-2)l}, x_{k,t+(d-1)l}\}$ are associated with numbers from 1 to d and arranged in increasing order as $\{x_{k,t+(j_1-1)l}, x_{k,t+(j_2-1)l}, ..., x_{k,t+(j_{d-1}-1)l}, x_{k,t+(j_d-1)l}\}$ [9]. For different samples, there will be d! potential ordinal patterns, π, named "motifs". For each π_t, $p(\pi_{k,t})$ contains the relative frequency as follows:

$$p(\pi_{k,i}^{d,l}) = \frac{\#\left|t : t \le N-d, \text{type}(X_{k,t}^{d,l}) = \pi_{k,i}^{d,l}\right|}{(N-d+1)P} \tag{2}$$

where # means cardinality [7, 14].

The differences between Eq. (2) and the correponding equation in the original definition of PE [7] is that the relative frequency in mvPE is devided by the number of channels P so that $\sum_{k=1}^{P}\sum_{j=1}^{d!} p_{k,j} = 1$ holds.

The marginal relative frequencies demonstrating the distribution of the motifs are defined as follows:

$$p_j = \sum_{k=1}^{P} p_{k,j} \tag{3}$$

Consequently, mvPE for each scale β, is deifind as:

$$mvPE(d,l,\mathbf{X}) = -\sum_{\pi_j=1}^{\pi_j=d!} p(\pi_j)\ln p(\pi_j) \tag{4}$$

When all marginal relative frequencies have equal probabilities, the largest value of PE is obtained, which has a value of $\ln(d!)$. In contrast, if there is only one $p(\pi_j)$ different from zero, which demonstrates a completely regular signal in every channel, the smallest value of PE is bounded by 0.

2) Multivariate Sample Entropy

To calculate the mvSE, multivariate embedded vectors are initially generated [5]. In [15], the Takens embedding theorem for multivariate concept is described. Assume we have a P-channel signal $\mathbf{X} = \{x_{k,i}\}_{k=1,\ i=1}^{k=P,\ i=N}$ where N is the length of each time series $\{\mathbf{x}_k\}_{k=1}^{P}$. The multivariate embedded reconstruction is defined as:

$$X_m(i) = [x_{1,i}, x_{1,i+\tau_1}, ..., x_{1,i+(m_1-1)\tau_1}, x_{2,i}, x_{2,i+\tau_2}, ..., x_{2,i+(m_2-1)\tau_2}, ... \\ x_{P,i}, x_{P,i+\tau_P}, ..., x_{P,i+(m_P-1)\tau_P}] \tag{5}$$

where $\mathbf{M} = [m_1, m_2, ..., m_p]$ and $\boldsymbol{\tau} = [\tau_1, \tau_2, ..., \tau_p]$ are the embedding and the time lag vectors, respectively.

For P-variate time series $\{\mathbf{x}_k\}_{k=1}^{P}$, the mvSE algorithm, as a natural extension of standard univariate sample entropy, is shown as follows [5]:

1. Form multivariate embedded vectors $X_m(i) \in R^m$ where i =1,2,...,N-n and n=max{$\mathbf{M}$}×max{$\boldsymbol{\tau}$}.
2. Calculate the distance between any two composite delay vectors $X_m(i)$ and $X_m(j)$ as the maximum norm.
3. For a given $X_m(i)$ and a threshold r, count the number of instances P_i where $d[X_m(i), X_m(j)] \le r,\ i \ne j$. Next, calculate the frequency of occurrence as $B_i^m(r) = \frac{1}{N-n}P_i$ and define a global quantity $B^m(r) = \frac{1}{N-n}\sum_{i=1}^{N-n} B_i^m(r)$.
4. Extend the dimensionality of the multivariate delay vector in (5) from m to (m+1) (keep the dimension of the other variables unchanged).
5. Repeat steps 1 to 4 and find $B_i^{(m_k+1)}(r)$. Next, calculate $B_i^{(m+1)}(r)$ which denotes the avearge over all k of $B_i^{(m_k+1)}(r)$. Finally, find $B^{(m+1)}(r)$ which stands for the avearge over all i of $B_i(r)$ in an (m+1)-dimensional space.
6. Finally, MVSE is defined as:

$$MVSE(\mathbf{M},\boldsymbol{\tau},r,N) = -\ln\left(\frac{B^{(m+1)}(r)}{B^{m}(r)}\right)$$

where r and N are respectively the tolerance level and the length of the time series, and $\mathbf{M}$ and $\boldsymbol{\tau}$ were defined earlier.

It should be added that because multi-channel signals may have different amplitude ranges, the distances calculated on embedded vectors may be biased toward the largest amplitude ranges variates. For this reason, we scale all the data channels to the same amplitude range, and select the range [0,1] as a preferred choice [5].

C. Multivariate Multiscale Permutation and Sample Entropies

To calculate the mvMPE or mvMSE, for each scale factor β, respectively, mvPE or mvSE of the coarse-grained time series are calculated.

Choosing an acceptable embedding dimension d in PE, mvPE and mvMPE is challenging. To work with reliable statistics when calculating PE, it is highly recommended $d! << N$ [14]. In addition, when d is too large, the computation time will be higher. While d is high, the number of accessible states will be large, and the value of the PE will probably be more reliable. All in all, we should make a trade-off between the aforementioned cases. For mvPE, because the number of samples increases to PN, an appropriate embedding dimension should follow $d! << PN$. Accordingly for mvMPE, since using scale factor β causes the length of signal decrease to $\frac{PN}{\beta}$, it is recommended $d! << \frac{PN}{\beta}$.

IV. Results and Discussion

To evaluate the abilities of mvMPE and mvMSE approaches using MEG signals in terms of regions, according to Fig. 3 in [16], five scalp areas (anterior, left and right lateral, central, and posterior) were defined. Both the mvMSE and mvMPE algorithms were applied for channels 15, 17, 19, 21, 23, 25, 27, 29 (central region), 31, 32, 48, 49, 51, 52, 69, 70 (anterior region), 64, 66, 68, 107, 111, 125, 127, 129 (right lateral region) 53, 55, 57, 96, 100, 114, 116, 118 (left lateral region), 39, 41, 59, 62, 102, 105, 120 and 123 (posterior region) with a maximum of scale factor $\beta = 10$, embedding dimension of 4 for mvMPE, embedding dimension of 1 for mvMSE and $r = 0.15 \times$(standard deviation of the normalized time series) for each data channel according to [6] and [5].

The mvPE and mvSE values of each coarse-grained sequence versus the scales are shown in Fig. 1 and Fig. 2, in that order. The error bars at each scale show the SD of the average of results for AD and control groups, which are illustrated in red and blue, respectively.

As can be seen in Fig. 1, the averages of mvPE values of AD patients in scale factor 1 are larger than those of the controls. This appears to contradict [3, 4] showing AD patients are less complex and more regular than in controls. However, in the deeper scale factors, for each region, either the average of mvPEs of the AD group is lower than that of controls or the differences between the results of two groups are small. This declares the importance of mvMPE and shows that the MEG contains information in deeper scales as well as the smallest one.

For anterior (A), central (C), and posterior (P) regions, the average values of mvMSE for controls are higher than (or approximately equal to) those of the corresponding AD patients. For right lateral (R) and left lateral (L) regions, in scale factors 1 to 6, the mvMSE values for AD patients are lower than those of the corresponding controls. However, for the regions R and L and scale factors 7 to 10, the averages of mvSE values of AD groups are larger than those of controls. This illustrates that, for these two regions and for scale factors 7 to 10, mvMSE values contradict what claimed in [3, 4]. However, there are no significant differences between groups.

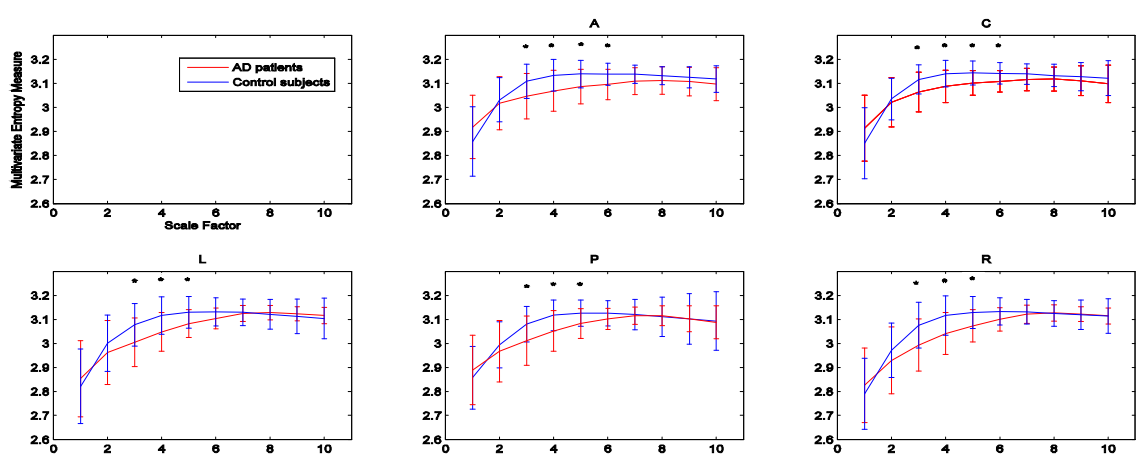

Fig. 1. Average values for mvMPE over 10 scale factors for AD (red) and control groups (blue) in each scalp region: anterior (A), central (C), right lateral (R), left lateral (L), and posterior (P). Bars indicate standard deviation. Asterisks indicate scales with significant differences between groups.

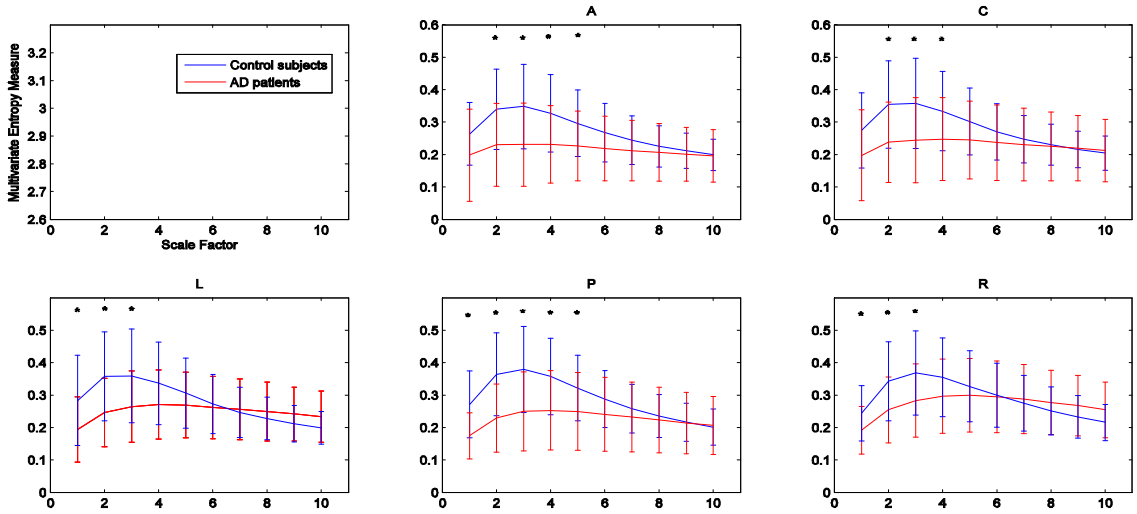

Fig. 2. Average values for mvMSE (b) over 10 scale factors for AD (red) and control groups (blue) for 5 scalp regions, described in Fig. 1.

A paired *t*-test was also run for AD patients vs. controls. We adjusted the false discovery rate independently for each multivariate entropy measure. The scales and regions were the adjusted *p*-values were significant are shown with * in Fig. 1 and Fig. 2. This shows that mvMPE, in all five regions, achieves significant differences at scales 3, 4, and 5. On the other hand, the behavior of mvMSE is less uniform across regions. Only scales 2 and 3 show significant differences in all regions.

In brief, we have studied the MEG background activity of AD patients using mvMPE and mvMSE. This study benefits from the consideration of MEG signals, which may have some advantages over EEGs, a larger sample size, and the consideration of five regions and more channels over the scalp, in comparison with [2]. In addition, we have discussed a proper embedding dimension for mvMPE and mvPE for the first time. The results showed that most significant differences appeared for $\beta>1$, thus highlighting the importance of multiscale evaluations of brain activity.

In future work, we intend to consider multivariate multiscale fuzzy entropy and parameterize the curves obtained by each technique to assess the performance of existing and novel metrics in the context of AD diagnosis. We will also investigate the effect of gender on the results.

REFERENCES

[1] J. R. Petrella, R. E. Coleman, and P. M. Doraiswamy, "Neuroimaging and Early Diagnosis of Alzheimer Disease: A Look to the Future 1," *Radiology,* vol. 226, pp. 315-336, 2003.

[2] D. Labate, F. La Foresta, G. Morabito, I. Palamara, and F. C. Morabito, "Entropic Measures of EEG Complexity in Alzheimer's Disease Through a Multivariate Multiscale Approach," *Sensors Journal, IEEE,* vol. 13, pp. 3284-3292, 2013.

[3] R. Hornero, D. Abásolo, J. Escudero, and C. Gómez, "Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences,* vol. 367, pp. 317-336, 2009.

[4] K. López-de-Ipina, J. Solé-Casals, H. Eguiraun, J. B. Alonso, C. M. Travieso, A. Ezeiza, *et al.*, "Feature selection for spontaneous speech analysis to aid in Alzheimer's disease diagnosis: A fractal dimension approach," *Computer Speech & Language,* vol. 30, pp. 43-60, 2015.

[5] M. U. Ahmed and D. P. Mandic, "Multivariate multiscale entropy: A tool for complexity analysis of multichannel data," *Physical Review E,* vol. 84, p. 061918, 2011.

[6] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology,* vol. 278, pp. H2039-H2049, 2000.

[7] C. Bandt and B. Pompe, "Permutation Entropy: A Natural Complexity Measure for Time Series," *Physical Review Letters,* vol. 88, pp. 1-4, 2002.

[8] M. Costa, A. L. Goldberger, and C.-K. Peng, "Multiscale Entropy Analysis of Complex Physiologic Time Series," *Physical Review Letters,* vol. 89, pp. 1-4, 2002.

[9] F. C. Morabito, D. Labate, F. La Foresta, A. Bramanti, G. Morabito, and I. Palamara, "Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG," *Entropy,* vol. 14, pp. 1186-1202, 2012.

[10] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain," *Reviews of modern Physics,* vol. 65, p. 413, 1993.

[11] J. Escudero, S. Sanei, D. Jarchi, D. Abásolo, and R. Hornero, "Regional coherence evaluation in mild cognitive impairment and Alzheimer's disease based on adaptively extracted magnetoencephalogram rhythms," *Physiological measurement,* vol. 32, p. 1163, 2011.

[12] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of Alzheimer's disease Report of the NINCDS-ADRDA Work Group* under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease," *Neurology,* vol. 34, pp. 939-939, 1984.

[13] J. Escudero, R. Hornero, D. Abásolo, and A. Fernández, "Quantitative evaluation of artifact removal in real magnetoencephalogram signals with blind source separation," *Annals of biomedical engineering,* vol. 39, pp. 2274-2286, 2011.

[14] M. Zanin, L. Zunino, O. A. Rosso, and D. Papo, "Permutation Entropy and Its Main Biomedical and Econophysics Applications: A Review," *Entropy,* vol. 14, pp. 1553-1577, 2012.

[15] L. Cao, A. Mees, and K. Judd, "Dynamics from multivariate time series," *Physica D: Nonlinear Phenomena,* vol. 121, pp. 75-88, 1998.

[16] J. Poza, R. Hornero, J. Escudero, A. Fernández, and C. Sánchez, "Regional Analysis of Spontaneous MEG Rhythms in Patients with Alzheimer's Disease Using Spectral Entropies," *Annals of Biomedical Engineering,* vol. 36, pp. 141-152, 2008/01/01 2008.