

RESEARCH ARTICLE

Quantile graphs for EEG-based diagnosis of Alzheimer's disease

Aruane M. Pineda¹, Fernando M. Ramos², Luiz Eduardo Betting³, Andriana S. L. O. Campanharo^{4*}

1 Department of Biostatistics, Institute of Biosciences, São Paulo State University (UNESP), Botucatu, São Paulo, Brazil, **2** National Institute for Space Research (INPE), Earth System Science Center (CCST), São José dos Campos, São Paulo, Brazil, **3** Department of Neurology, Psychology and Psychiatry, Institute of Biosciences, Botucatu Medical School, São Paulo State University (UNESP), Botucatu, São Paulo, Brazil, **4** Department of Biostatistics, Institute of Biosciences, São Paulo State University (UNESP), Botucatu, São Paulo, Brazil

* andriana.campanharo@unesp.br

OPEN ACCESS

Citation: Pineda AM, Ramos FM, Betting LE, Campanharo ASLO (2020) Quantile graphs for EEG-based diagnosis of Alzheimer's disease. PLoS ONE 15(6): e0231169. <https://doi.org/10.1371/journal.pone.0231169>

Editor: Ruxandra Stoean, University of Craiova, ROMANIA

Received: December 2, 2019

Accepted: March 17, 2020

Published: June 5, 2020

Copyright: © 2020 Pineda et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The database was designed jointly by researchers at Florida State University and made freely available by Campanharo et al. It can be freely downloaded at <https://osf.io/s74qf/>, as mentioned in the paper.

Funding: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. L. E. B. acknowledges the support of São Paulo Research Foundation (FAPESP), grant 2016/17914-3. A. S. L. O. C. acknowledges the support of São Paulo Research Foundation (FAPESP),

Abstract

Known as a degenerative and progressive dementia, Alzheimer's disease (AD) affects about 25 million elderly people around the world. This illness results in a decrease in the productivity of people and places limits on their daily lives. Electroencephalography (EEG), in which the electrical brain activity is recorded in the form of time series and analyzed using signal processing techniques, is a well-known neurophysiological AD biomarker. EEG is noninvasive, low-cost, has a high temporal resolution, and provides valuable information about brain dynamics in AD. Here, we present an original approach based on the use of quantile graphs (QGs) for classifying EEG data. QGs map frequency, amplitude, and correlation characteristics of a time series (such as the EEG data of an AD patient) into the topological features of a network. The five topological network metrics used here—clustering coefficient, mean jump length, betweenness centrality, modularity, and Laplacian Estrada index—showed that the QG model can distinguish healthy subjects from AD patients, with open or closed eyes. The QG method also indicates which channels (corresponding to 19 different locations on the patients' scalp) provide the best discriminating power. Furthermore, the joint analysis of delta, theta, alpha, and beta wave results indicate that all AD patients under study display clear symptoms of the disease and may have it in its late stage, a diagnosis known a priori and supported by our study. Results presented here attest to the usefulness of the QG method in analyzing complex, nonlinear signals such as those generated from AD patients by EEGs.

1 Introduction

Alzheimer's disease (AD) is the main cause of dementia in people over 65 years of age, affecting nearly 25 million people throughout the world [1]. AD is marked primarily by progressive cognitive impairment, loss of memory, and disorientation of time and space. [2]. With an unknown cause, AD usually evolves slowly, following a specific pathway that first involves the

grants 2018/25358-9 and 2020/04989-0. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

hippocampus, then spreads out to association areas in parietal, lateral temporal, and frontal regions, eventually affecting all regions of cortex [3, 4]. AD is irreversible. Thus, the earlier a diagnosis can be made, the sooner the treatment can be started, with a higher chance of success in slowing down the progression of the disease.

Currently, definitive diagnosis of AD is made on the examination of the brain tissue accessed by biopsy or necropsy [5]. Since it is possible to be sure that he or she had AD only after the patient's death, clinical diagnosis is made by excluding other causes of dementia. However, in recent years there has been considerable research toward the diagnosis of AD using biological markers (biomarkers); see [6–9] for a review. Classified into four basic groups, biochemistry, genetics, neurophysiology, and neuroimaging, AD biomarkers should carry enough information on pathophysiologic processes active in AD so as to allow the detection of the disease's precursor clues before the symptoms onset [7].

Electroencephalography (EEG), in which the electrical brain activity is recorded in the form of time series and analyzed using signal processing techniques, is a well-known neurophysiological AD biomarker. EEG is noninvasive, low-cost, has a high temporal resolution, and provides valuable information about brain dynamics in AD [10–15]. Recently, computer-aided classification methods have been developed and applied to EEG signals to distinguish among patients with AD, healthy controls, and patients with mild cognitive impairment [9]. AD affects the characteristics of EEGs. Consequently, EEG analysis can provide useful information about the dynamics of the brain due to AD. Slowing EEG, decreased EEG coherence, and decreased EEG complexity are the most distinctive traits in the EEG caused by AD [16]. Nonetheless, the diagnosis of AD from EEG data is still open research topic, and it comes without surprise the wealth of methods proposed in the medical and related literature. These methods, based on Fast Fourier Transform (FFT) [17–19], Wavelet Transform (WT) [20–23], Phase-Space Reconstruction [24–26], Eigenvector Methods (EMs) [27, 28], Time Frequency Distributions (TFDs) [29], and the Auto-Regressive Method (ARM) [30], generally require from the input signal one or more of the following assumptions: stationarity, high time or frequency resolution, and/or a high signal-to-noise ratio.

In [31], we introduced a simple method for transforming a time series into graphs or networks, called the quantile graph (QG) method. The QG method maps relevant properties of the original time series, such as periodicity or randomness, into the topological features of the resulting graph. The QG method has been applied to quantify the Hurst exponent of gaussian white noises and brownian motions [32] and to uncover distinctions between physiological signals of normal individuals and unhealthy patients [13, 31]. We have also used the QG method to distinguish in an EEG the pre-ictal from the ictal stage of an epileptic convulsion [14]. Along the same lines, here we study the QG method as a technique to differentiate patients with AD from control using EEG signals and to indicate which channels (corresponding to 19 different locations on the patients' scalp) provide the best discriminating power.

After this introduction, Section 2 and 3 present, respectively, the QG method and the network measures used here for the characterization of complex networks. The EEGs used in this research are described in Section 4 while the corresponding results are examined in Section 5. Finally, Section 6 presents the pertinent conclusions.

2 Materials and methods

As described in detail previously [14, 31–33], the QG method converts a time series $X = \{x(t) | t \in \mathbb{N}, x(t) \in \mathbb{R}\}$ into a complex network $g = \{\mathcal{N}, \mathcal{A}\} \in \mathcal{G}$, with \mathcal{N} vertices (or nodes) and \mathcal{A} edges (or arcs). In the QG method, each quantile q_i for $i = 1, 2, \dots, Q$ of X is attributed to a node $n_i \in \mathcal{N}$ in g . Two nodes n_i and n_j are connected with a weighted arc $(n_i, n_j, w_{ij}^k) \in \mathcal{A}$

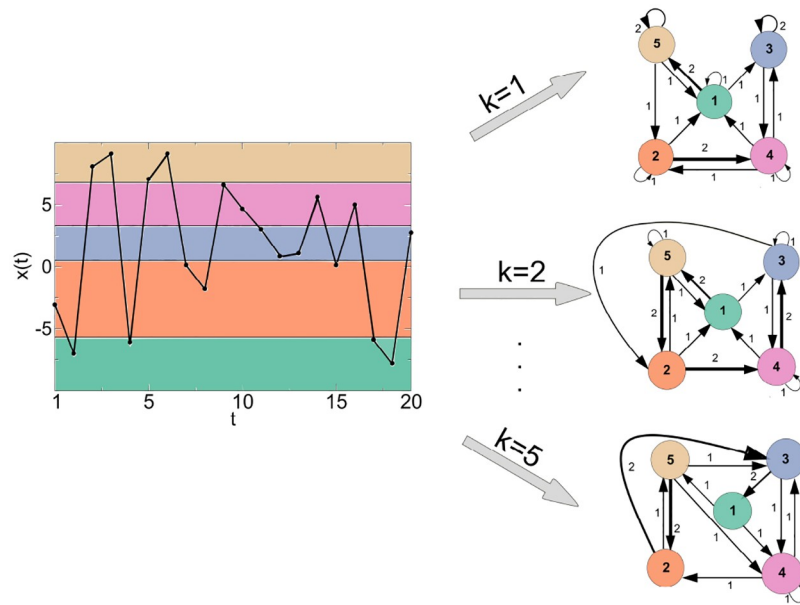


Fig 1. Example of the QG method for a time series with $T = 20$, $Q = 5$, and $k = 1, 2$ and 5 . The quantile intervals for the sorted data are given by $[x(0), x(4)[$, $[x(4), x(8)[$, $[x(8), x(12)[$, $[x(12), x(16)[$, and $[x(16), x(20)]$, i.e., $[-7.783, -3.050 [$, $[-3.050, 0.829[$, $[0.829, 4.657[$, $[4.657, 7.070[$, and $[7.070, 9.090]$. The quantiles are mapped into three networks with $\mathcal{N} = 5$ nodes each and arc weights given by $(1,1,1)$, $(1,3,1)$, $(1,5,2)$, $(2,1,1)$, $(2,2,1)$, $(2,4,2)$, $(3,3,2)$, $(3,4,1)$, $(4,1,1)$, $(4,2,1)$, $(4,3,1)$, $(4,4,1)$, $(5,1,1)$, $(5,2,1)$, $(5,5,2)$ for $k = 1$; $(1,3,1)$, $(1,5,2)$, $(2,1,1)$, $(2,4,2)$, $(2,5,1)$, $(3,2,1)$, $(3,3,1)$, $(3,4,1)$, $(4,1,1)$, $(4,3,2)$, $(4,4,1)$, $(5,1,1)$, $(5,2,2)$, $(5,5,1)$ for $k = 2$; and $(1,4,1)$, $(1,5,1)$, $(2,3,2)$, $(2,5,1)$, $(3,1,2)$, $(3,4,1)$, $(4,2,1)$, $(4,3,1)$, $(4,4,1)$, $(5,2,2)$, $(5,3,1)$, $(5,4,1)$ for $k = 5$.

<https://doi.org/10.1371/journal.pone.0231169.g001>

whenever two values $x(t)$ and $x(t + k)$ belong respectively to quantiles q_i and q_j , with $t = 1, 2, \dots, T$ and $k = 1, \dots, k_{max} < T$. Weights w_{ij}^k in the weighted directed adjacency matrix, which is denoted as A_k , are equal to the number of times q_i at time t is followed by q_j at time $t + k$. Thus, repeated transitions through the same edge increase the the corresponding weight value [14]. Normalizing A_k , it becomes a Markov transition matrix W_k with $\sum_j^Q w_{ij}^k = 1$ [32]. As shown in [14, 32], the QG method is weakly dependent on the choice of Q . Here, the number of quantiles is given by $Q \approx 2T^{1/3}$ [34]. A C-code implementation of our method has been made freely available by Pineda *et al.* [35].

As an illustrative example, Fig 1 shows a simple application of the QG method, X with $T = 20$ and $Q = 5$ (colored shading). Quantiles are defined as $[x(0), x(4)[$, $[x(4), x(8)[$, $[x(8), x(12)[$, $[x(12), x(16)[$, and $[x(16), x(20)]$. For a given k , the quantiles are mapped into a network with $\mathcal{N} = 5$ nodes and each quantile is assigned to a node $n_i \in \mathcal{N}$ in the corresponding network g . For $k = 1, 2$, and 5 , the nodes n_i and n_j are connected with weighted arcs (n_i, n_j, w_{ij}^1) , (n_i, n_j, w_{ij}^2) , and $(n_i, n_j, w_{ij}^5) \in \mathcal{A}$, respectively. The arc weights are given by $(1,1,1)$, $(1,3,1)$, $(1,5,2)$, $(2,1,1)$, $(2,2,1)$, $(2,4,2)$, $(3,3,2)$, $(3,4,1)$, $(4,1,1)$, $(4,2,1)$, $(4,3,1)$, $(4,4,1)$, $(5,1,1)$, $(5,2,1)$, $(5,5,2)$ for $k = 1$; $(1,3,1)$, $(1,5,2)$, $(2,1,1)$, $(2,4,2)$, $(2,5,1)$, $(3,2,1)$, $(3,3,1)$, $(3,4,1)$, $(4,1,1)$, $(4,3,2)$, $(4,4,1)$, $(5,1,1)$, $(5,2,2)$, $(5,5,1)$ for $k = 2$; and $(1,4,1)$, $(1,5,1)$, $(2,3,2)$, $(2,5,1)$, $(3,1,2)$, $(3,4,1)$, $(4,2,1)$, $(4,3,1)$, $(4,4,1)$, $(5,2,2)$, $(5,3,1)$, $(5,4,1)$ for $k = 5$. Note that the more repeated transitions between the quantiles q_i and q_j occur, the larger the weights between the nodes n_i and n_j are (represented in the corresponding network by thicker lines).

3 Network measures

In the recent past, several studies have shown the relevance and usefulness of complex network theory to the comprehension of a wide range of phenomena, across various scientific disciplines, from social sciences to biology [36]. Complex network theory relies on the use of mathematical metrics that can to quantify different features of the network's topology. Based on the adjacency matrix A and Markov transition matrix W , we describe the network measures used in this work, namely the clustering coefficient (CC), the mean jump length (Δ), the betweenness centrality (BC), the modularity (Mo), and the Laplacian Estrada index (LEE). Code implementations of those measures have been made freely available by Pineda *et al.* [35] and Bounova [37].

3.1 Clustering coefficient

Some networks tend to have more links between adjacent vertices, in a way that their topology deviates from that of an uncorrelated random network, in which triangles are sparse. This pattern is called clustering [38], and reflects the segregation of edges into tightly connected neighborhoods. There have been various attempts in the literature to develop a clustering coefficient for weighted networks. Here, the clustering coefficient of a given node n_i is given by [38]:

$$CC_i = \frac{1}{s_i(d_i - 1)} \frac{\sum_{j,d} (w_{ij} + w_{id})}{2} (a_{ij}a_{jd}a_{id}), \quad (1)$$

where w_{ij} is an element in the weighted matrix W , $a_{ij} = 1$ if there is a arc from node n_i to node n_j , and 0 otherwise. d_i is the total degree of node n_i , and s_i is the strength of connectivity of node n_i . The global clustering coefficient for the entire network, denoted as CC , is defined by the average of the local clustering coefficients over all nodes.

3.2 Mean jump length

Given a Markov transition matrix W of a graph g , it is possible to perform a random walk on it and compute the mean jump length Δ , defined as follows [32]:

$$\Delta = \frac{1}{S} \sum_{s=1}^S \delta_s(i, j), \quad (2)$$

where $s = S$ is the total number of jumps, and the length $\delta_s(i, j) = |i - j|$, with $i, j = 1, \dots, Q$ being the node indices, as defined by W . As described previously [14], a less time-consuming approach for the calculation of Δ , for large S , is given by:

$$\Delta = \frac{1}{Q} \text{tr}(PW^T), \quad (3)$$

with W^T being the transpose of W , P a $Q \times Q$ matrix with elements $p_{ij} = |i - j|$, and tr the trace operation.

3.3 Betweenness centrality

Betweenness is a centrality measure based on shortest paths, widely used in complex network analysis. The betweenness centrality (BC) of a node n_u is given by [39]:

$$BC_{n_u} = \sum_{ij} \frac{\sigma(n_i, n_u, n_j)}{\sigma(n_i, n_j)}, \quad (4)$$

where $\sigma(n_i, n_w, n_j)$ is the number of shortest paths between nodes n_i and n_j that go through node n_w , $\sigma(n_i, n_j)$ is the total number of shortest paths between n_i and n_j , and the sum is calculated over all pairs n_i, n_j of distinct nodes [39, 40]. The betweenness centrality for the entire network, denoted by BC , is defined as the average of the local betweenness centralities over all nodes.

3.4 Modularity

Recently, the subject of detecting the modular structure of a complex network has gained a large amount of attention [41]. Networks with high modularity present smaller clusters of nodes connected more to each other than to the network at large [42]. Several methodologies have been developed for modules detection and characterization [41]. The goal of a module identification algorithm is to find P_i that maximizes the modularity $M(P_i)$, where P_i is the set of nodes of module i . Given the ensemble \mathcal{P} of all partitions, the modularity of $P \in \mathcal{P}$ is computed as:

$$M(P) = \sum_{i=1}^m \left[\frac{e_i}{E} - \left(\frac{d_i}{2E} \right)^2 \right], \quad (5)$$

with E being the total number of edges in the network, d_i the sum of all node degrees in module i , and e_i the number of edges within module i . In Eq (5), the sum is evaluated over all the m modules in the partition P [43]. In the present work, we used the algorithm developed by [44] for determining \mathcal{P} and calculating $M(P)$.

3.5 Laplacian estrada index

Let g be a network without loops and multiple edges. The Laplacian matrix of g is the matrix $L = D - A$ where D is a diagonal matrix with (d_1, \dots, d_n) on the main diagonal in which d_i is the degree of the node n_i . Since L is a real symmetric matrix, its eigenvalues $\mu_1, \mu_2, \dots, \mu_n$ are real numbers. These are referred to as the Laplacian eigenvalues of the underlying network [45]. Let's assume those to be labelled in a non-increasing manner $\mu_1 \geq \mu_2 \dots \geq \mu_n$. The Laplacian Estrada index of a network g is defined as [46]:

$$LEE = LEE(g) = \sum_{i=1}^n e^{\mu_i}. \quad (6)$$

4 Data

The database was designed jointly by researchers at Florida State University and it was recorded from the 19 scalp ($F_{p1}, F_{p2}, F_z, F_3, F_4, F_7, F_8, C_z, C_3, C_4, T_3, T_4, P_z, P_3, P_4, T_5, T_6, O_1,$ and O_2) loci of the international 10-20 system using a Biologic Systems Brain Atlas III Plus workstation [47]. The letters F, C, P, O, and T refer to cerebral lobes (F: frontal, C: central, P: parietal, O: occipital, and T: temporal). Recordings, which include four groups (denoted as A, B, C, and D), were made under two rest states: eyes open (groups A and C) by visually fixating and eyes closed (groups B and D) using a linked-mandible reference forehead ground [47]. Groups A and B represent healthy controls and consist of 24 healthy elderly (average age 72, range 61-83), all being negative for any neurological or psychiatric disorders. Groups C and D consist of 24 probable AD patients (average age 69, range 53-85) diagnosed through the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) and the Alzheimer's Disease and Related Disorders Association (ADRDA), and Diagnostic and

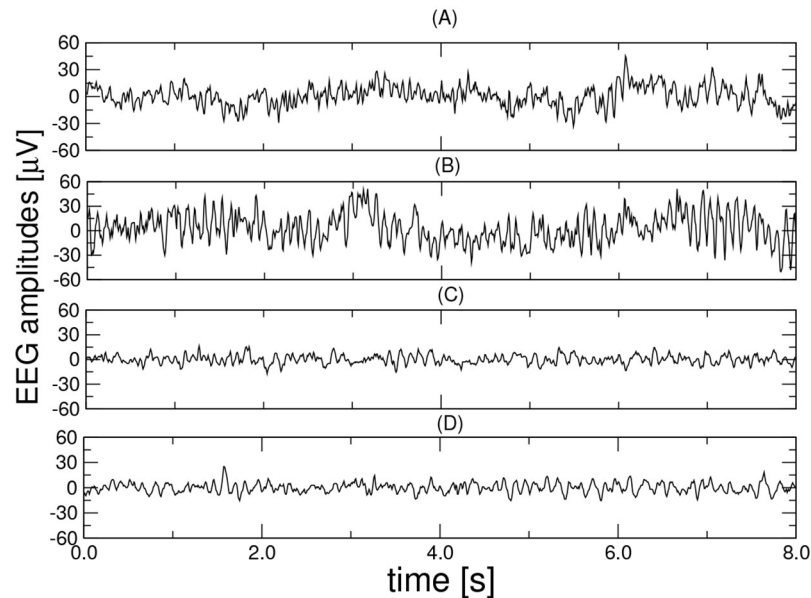


Fig 2. Exemplary EEG segments (channel F_7) from each of the four groups (A, B, C, and D). From top to bottom: health controls, eyes open (group A), health controls, eyes closed (group B), patient with AD, eyes open (group C) and patient with AD, eyes closed (group D).

<https://doi.org/10.1371/journal.pone.0231169.g002>

Statistical Manual of Mental Disorders (DSM)-III-R criteria, as described previously [33, 48]. EEG segments of 8 s duration, band-limited to the range of 1-30 Hz, were recorded at a sampling frequency of 128 Hz (free from eye motion and blinking, and myogenic artifacts) and extracted from the EEG recordings. An EEG technician was with each patient during the recordings to monitor the patients' vigilance state. The database has been provided by Dr. Dennis Duke of Florida State University and made freely available by Campanharo *et al.* [49]. A detailed description of the database can be found in [47]. Exemplary data from channel F_7 are depicted in Fig 2.

5 Results

5.1 Discriminating between aging and Alzheimer's disease

We apply the QG method to the problem of discriminating patients with AD from normal controls. The data from channel F_7 was chosen in the simulations due to its closeness to the hippocampal region, which is one of the first regions of the brain to be affected by AD. As all time series have the same length ($T = 1,024$), we used $Q = 2(1,024)^{1/3} \approx 20$ and $k = 1, 2, \dots, 25$ in all calculations. Thus, we mapped $2 \times 24 \times 25$ time series into 1,200 quantile graphs (or 1,200 A_k matrices), and therefore, we obtained 1,200 W_k matrices with $Q^2 = 400$ elements each. Following, for each group and a given k , we computed the median over all weighed directed adjacency matrices A_k and obtained the Markov transition matrix of medians. For all groups, we computed $CC(k)$, $\Delta(k)$, $BC(k)$, $Mo(k)$, and $LEE(k)$ versus k using Eqs (1), (3), (4), (5) and (6), respectively (Fig 3). Note in all cases that the curves for normal controls (A or B) and patients with AD (C or D) form two distinct clusters with maximum separation at approximately $k_{max} = 9$ for $CC(k)$, $k_{max} = 10$ for $\Delta(k)$, $k_{max} = 6$ for $BC(k)$, $k_{max} = 6$ for $Mo(k)$, and $k_{max} = 8$ for $LEE(k)$.

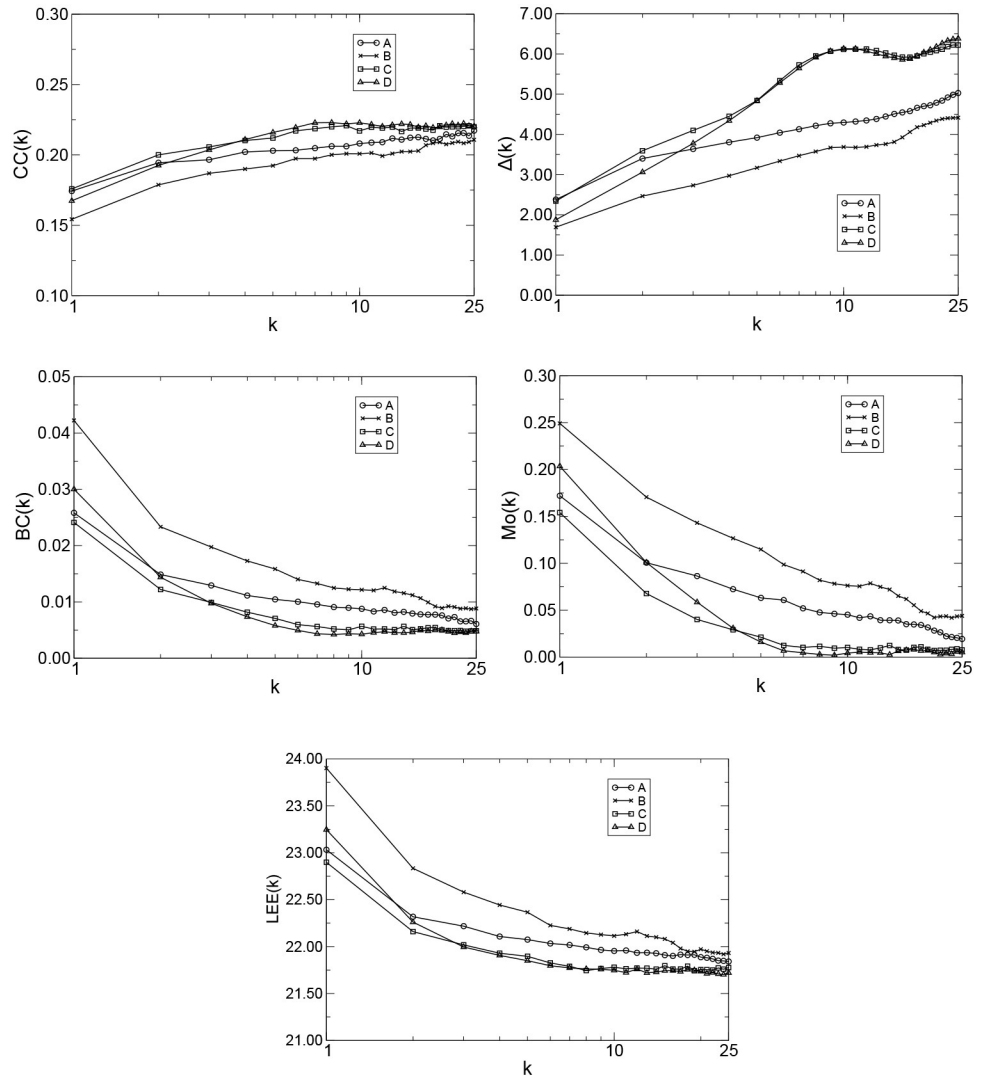


Fig 3. $CC(k)$, $\Delta(k)$, $BC(k)$, $Mo(k)$, and $LEE(k)$ versus k , $T = 1, 024$, $Q = 20$, and $k = 1, 2, \dots, 25$ for the groups A (patients from health controls, eyes open), B (patients from health controls, eyes closed), C (patients with AD, eyes open), and D (patients with AD, eyes closed).

<https://doi.org/10.1371/journal.pone.0231169.g003>

Fig 4 presents boxplots of $CC(k)$, $\Delta(k)$, $BC(k)$, $Mo(k)$, and $LEE(k)$, computed for twelve sample segments each, for the groups A, B, C, and D, and $k_{max} = 9$, $k_{max} = 10$, $k_{max} = 6$, $k_{max} = 6$, and $k_{max} = 8$, respectively. We observe that, irrespective of the metric used to characterize a given network, the QG method identified health controls with eyes open (group A) and closed (group B), and patients with AD with eyes open (group C) and closed (group D). We performed ANOVA analysis to quantify the sample mean differences found in Fig 4. Table 1 shows a 95% confidence interval and a p-value of less than 0.05 among the sample means of the network measures CC , Δ , BC , Mo , and LEE for the groups A, B, C, and D. We also performed a Receiver Operating Characteristic (ROC) analysis [50, 51] in order to quantify how accurately the QG was able to discriminate subjects and/or patients from any two groups under different health conditions. Table 2 shows the areas under the ROC curves (A_{ROC}) of the

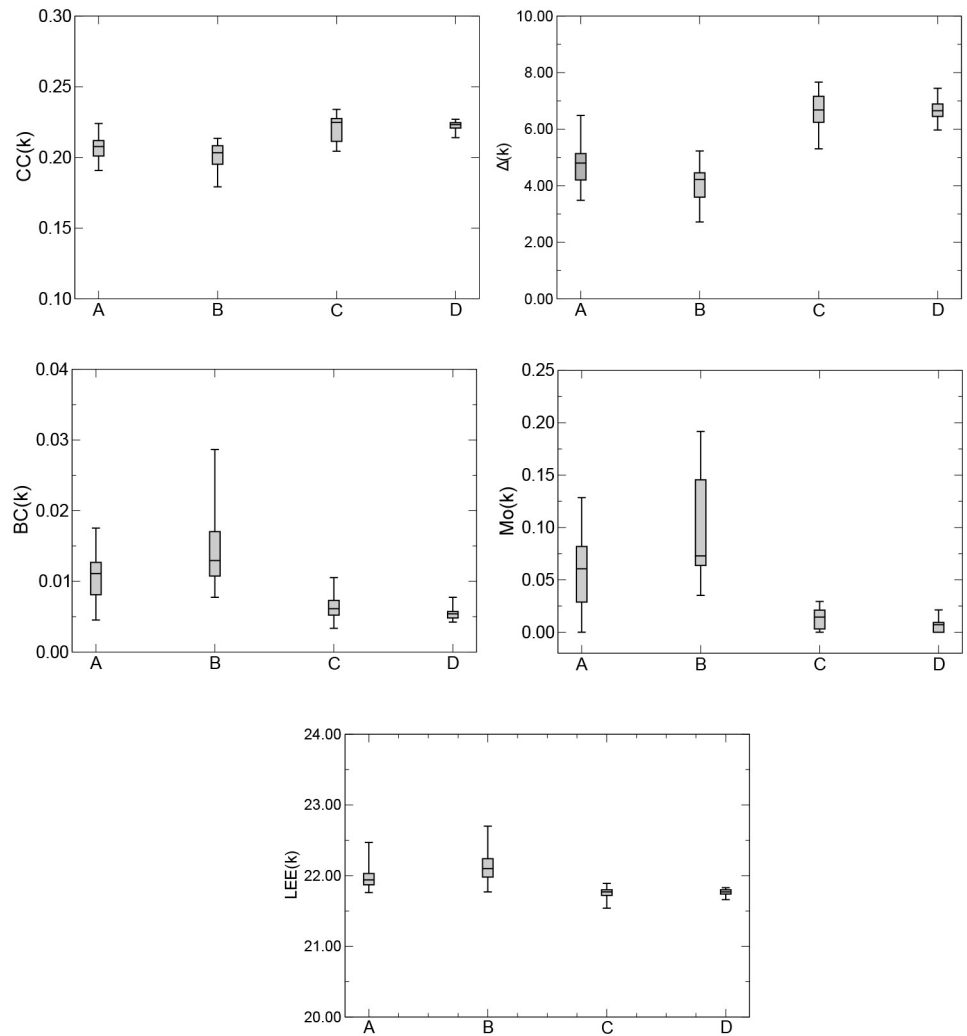


Fig 4. Boxplots of $CC(k)$, $\Delta(k)$, $BC(k)$, $Mo(k)$, and $LEE(k)$ for $k_{max} = 9$, $k_{max} = 10$, $k_{max} = 6$, $k_{max} = 6$, and $k_{max} = 8$, respectively, for the groups A, B, C, and D. Boxplots from patients with different health states show different means (placed at the center of each box), which are 0.2077, 0.2033, 0.2248, and 0.2232 for $CC(k)$; 4.8010, 4.2190, 6.6790, and 6.6510 for $\Delta(k)$; 0.0111, 0.0129, 0.0061, and 0.0054 for $BC(k)$; 0.0605, 0.0728, 0.0145, and 0.0072 for $Mo(k)$; and 21.9400, 22.1000, 21.7700, and 21.7700 for $LEE(k)$, respectively.

<https://doi.org/10.1371/journal.pone.0231169.g004>

Table 1. Statistical comparison (95% confidence interval, $p < 0.05$) among the sample means of the network measures CC, Δ, BC, Mo, and LEE for the groups A, B, C and, D, through ANOVA.

	CC	Δ	BC	Mo	LEE
CI_{HU}	[-0.0221; -0.0071]	[-2.5075; -1.3472]	[0.0012; 0.0078]	[0.0200; 0.0751]	[0.0978; 0.3967]
$p - value$	2.3900×10^{-8}	1.9000×10^{-14}	2.3360×10^{-7}	5.5930×10^{-9}	3.2160×10^{-7}

$U = \{C, D\}$ (unhealthy)

$H = \{A, B\}$ (healthy)

<https://doi.org/10.1371/journal.pone.0231169.t001>

Table 2. Areas under the ROC curves (A_{ROC}) of the network measures CC , Δ , BC , Mo , and LEE , for patients in groups A and C, B and D, A and D, and B and C for $k_{max} = 9, 10, 6, 6$ and 8 , respectively.

	CC	Δ	BC	Mo	LEE
$A_{ROC_{AC}}$	0.8472	0.9653	0.8368	0.8785	0.8889
$A_{ROC_{BD}}$	1.0000	1.0000	0.9965	1.0000	0.9583
$A_{ROC_{AD}}$	0.9375	0.9653	0.9167	0.9236	0.9236
$A_{ROC_{BC}}$	0.9167	1.0000	0.9792	1.0000	0.9514

<https://doi.org/10.1371/journal.pone.0231169.t002>

network metrics CC , Δ , BC , Mo , and LEE , for patients in groups A and C, B and D, A and D and B and C. In all cases, the QG method shows an excellent performance in discriminating patients with different health conditions. Comparing the results across the five network metrics, we also observe that Δ provides the best discriminating power.

Finally, we used a support vector machine (SVM, [18, 52]), which is a supervised machine for two-class classification problems, to individually classify healthy elderly subjects and patients with AD. Based on the values of $CC(k)$, $\Delta(k)$, $BC(k)$, $Mo(k)$, and $LEE(k)$ of 24 healthy subjects (groups A and B) and 24 patients with AD (groups C and D) and the k-fold cross-validation strategy ($k = 10$), the patients were randomly divided into ten equivalent subsamples. Among the ten subsamples, nine-fold (90% of samples) were considered the training set and the remaining fold (10% of samples) was considered the test set. The values of accuracy (ACC) (100%), sensitivity (SEN) (100%), specificity (SPE) (100%), and area under the curve (AUC) (1.0) show that the QG method is a reliable technique for differentiating patients in different health conditions.

5.2 EEG channels influence in the AD detection

To verify the extent to which the electrode location affects the differentiation between normal controls and patients with AD, we used the 19 EEG channels available in our analysis. Analogous to previous case, since for all channels the time series have the same length ($T = 1, 024$), we used $Q = 2(1, 024)^{1/3} \approx 20$ and $k = 1, 2, \dots, 25$ in all calculations. Therefore, we mapped $19 \times 2 \times 24 \times 25$ time series into 22,800 quantile graphs (or 22,800 A_k matrices), and therefore, we obtained 22,800 W_k matrices with $Q^2 = 400$ elements each. For all groups and all channels, we calculated $CC(k)$, $\Delta(k)$, $BC(k)$, $Mo(k)$, and $LEE(k)$ versus k . For a given network measure and channel, k_{max} was chosen in such way to obtain the maximum separation between the curves of normal controls (groups A or B) and patients with AD (groups C or D) and the average of A_{ROC} , denoted here by \hat{A}_{ROC} , was computed through the combination between the groups.

Fig 5 shows the location on scalp of the 19 EEG channels, colored according to the value of \hat{A}_{ROC} for CC , Δ , BC , Mo , and LEE , respectively. The color map indicates values close to one for \hat{A}_{ROC} in all cases, which means that the QG method was effective in differentiating normal controls from patients with AD, regardless of the network measure and the electrode location. Overall, Δ is the metric that displays the best results with $0.9183 \leq \hat{A}_{ROC} \leq 1.000$. This result corroborates the knowledge that all patients under study display clear symptoms of the disease. Regardless of the network measure, the brain damage was found mostly in the parietal lobes and some loci in the temporal lobes (T_5 and T_6).

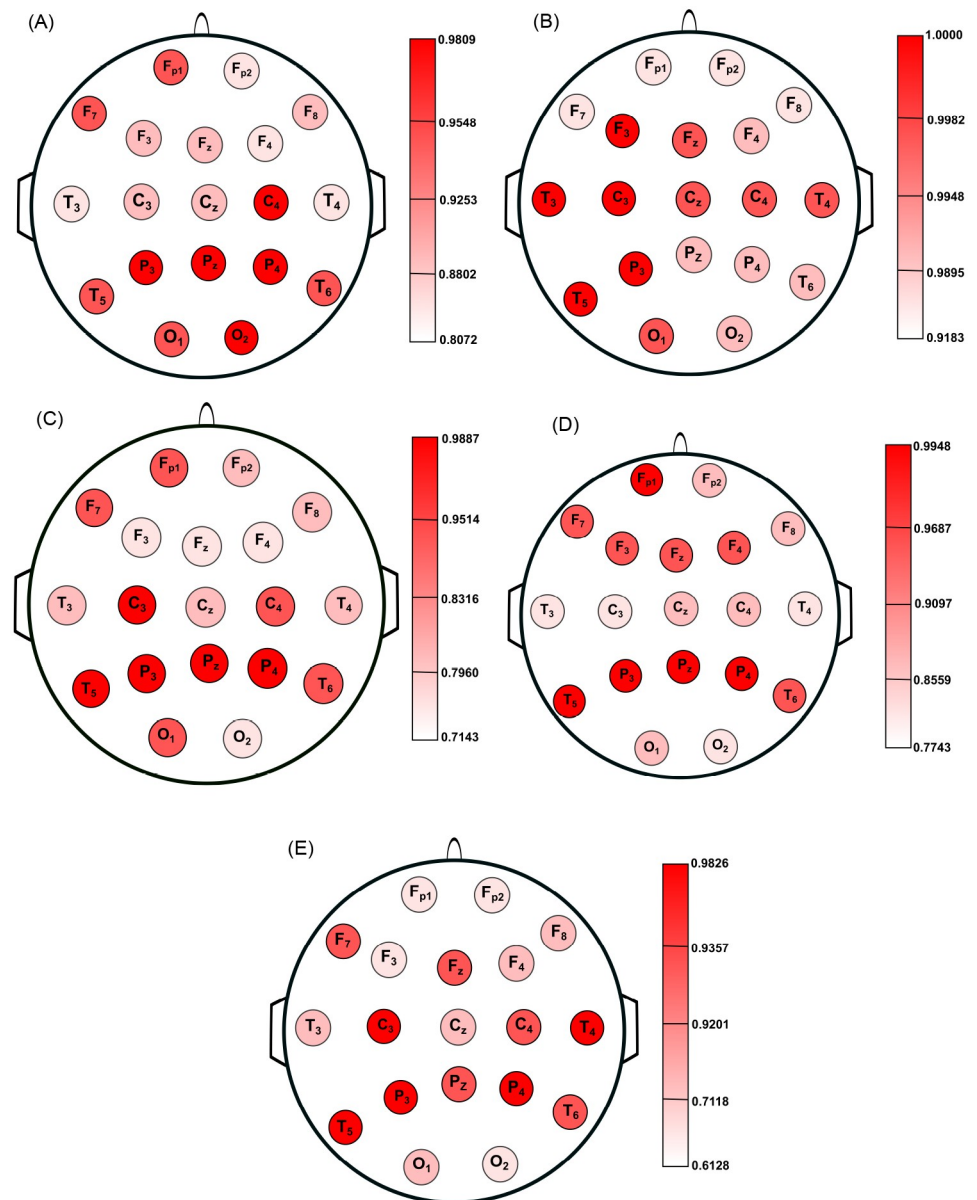


Fig 5. Location on scalp of the 19 EEG channels, represented by circles and colored according to the value of \hat{A}_{ROC} for CC (A), Δ (B), BC (C), Mo (D), and LEE (E), respectively. Circles with darker colors indicate a better differentiation between aging and AD.

<https://doi.org/10.1371/journal.pone.0231169.g005>

5.3 EEG wave patterns in the AD detection

It is widely accepted that Alzheimer's disease earliest changes are an increase in theta activity and a decrease in beta activity, which are followed by a decrease in alpha activity [53, 54]. Delta activity increases later during the course of the disease [53]. In particular, the increase in theta activity is a typical finding in mild AD. The increase in delta activity is not evident until the more advanced stages of the disease take place [16, 54–56]. We apply the QG method to check

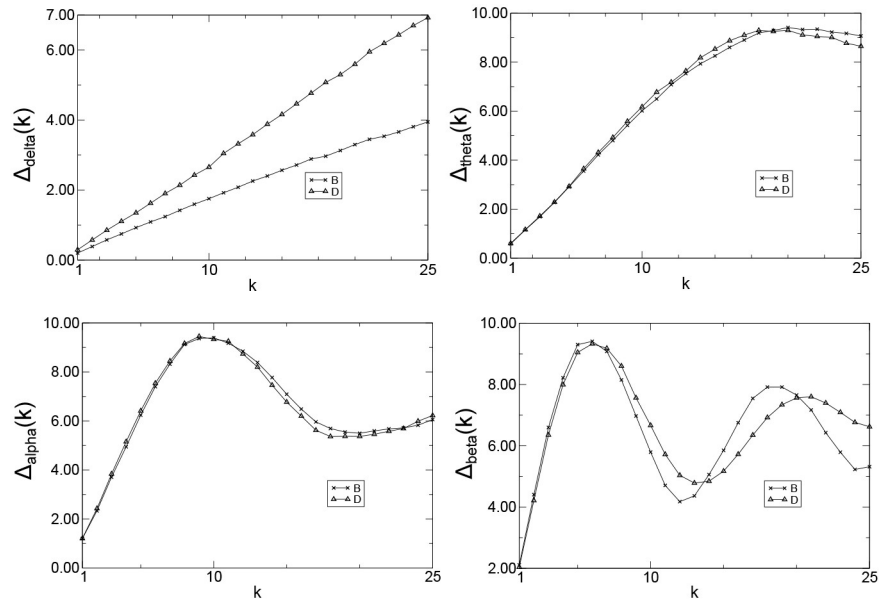


Fig 6. $\Delta(k)$ versus k (channel P_3), $T = 1, 024$, $Q = 20$, and $k = 1, 2, \dots, 25$ for delta (Δ_{delta}), theta (Δ_{theta}), alpha (Δ_{alpha}), and beta (Δ_{beta}) waves and patients for the groups B and D.

<https://doi.org/10.1371/journal.pone.0231169.g006>

if it can capture the influence of the EEG wave patterns in the AD development. Since the alpha rhythm increases in amplitude at rest with eyes closed [57], groups B and D were used in the analysis. Moreover, the data from channel P_3 was chosen in the simulations due to its closeness to the parietal region, which is one of the regions to be later affected by AD.

Wavelet digital filter [58, 59] were used to extract the four EEG frequency bands, i.e., delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), and beta (13-30 Hz). Analogous to previous cases, as all time series have the same length ($T = 1, 024$), we used $Q = 2(1, 024)^{1/3} \approx 20$ and $k = 1, 2, \dots, 25$ in all calculations. Thus, we mapped $4 \times 24 \times 25$ time series into 2,400 quantile graphs (or 2,400 A_k matrices), and therefore, we obtained 2,400 W_k matrices with $Q^2 = 400$ elements each. Following, for each group and wave and for a given k , we computed the median over all weighed directed adjacency matrices A_k and obtained the Markov transition matrix of medians.

For all groups and frequency bands, the mean jump length, which is the metric that best discriminates the groups in our analysis, was computed (Fig 6). Observe that the curves for normal controls (B) and patients with AD (D) are very similar for theta, alpha, and beta waves, regardless the value of k . On the other hand, there was statistically significant difference (95% confidence interval (CI) and a p -value of less than 0.05) between the sample means in groups B and D for delta waves. (Table 3). This result confirms the prior knowledge that all patients under study display clear symptoms of the disease and may have it in its late stage.

Table 3. Statistical comparison between the sample means of the network measure Δ for the groups B and D.

Waves	Delta	Theta	Alpha	Beta
CI_{BD}	[0.4701; 2.3893]	[-1.6573; 1.7063]	[-1.2420; 1.1546]	[-0.8414; 1.1609]
p - value	0.0045	0.9767	0.9418	0.7497

<https://doi.org/10.1371/journal.pone.0231169.t003>

6 Conclusion

Presently, there is no conclusive technique for the accurate diagnosis of AD [60, 61], a highly incapacitating disease. Thus, an automatic computer implemented technique based solely on the analysis of EEG data would potentially have a broad application. Building upon the work described in [33], in this study we presented an application of QG method to the analysis of EEG data. QGs map frequency, amplitude, and correlation characteristics of a time series (such as the EEG data of an AD patient) into the topological features of a network. The five network topological measures used here showed that the QG method is capable of discriminating health controls (with eyes open or closed) from patients with AD (with eyes open or closed), and indicate which channels (corresponding to 19 different locations on the patients' scalp) provide the best discriminating power. All five network topological measures were able to generate statistically robust positive AD diagnostics, although the mean jump length provided the best results. Moreover, the combination of the network measures with a machine learning technique achieved outstanding performance in the two-class pattern classification problem presented here.

Spatially, the electrodes that best captured the symptoms were those nearer to the left and temporal-parietal chains. This observation is in line with the current understanding of the AD progression. Generally, AD mainly affects the left side of the temporal-hippocampal network, which is responsible for verbal memory and, apparently, is a more vulnerable hemisphere [62]. Furthermore, the joint analysis of delta, theta, alpha, and beta wave results indicate that all AD patients under study display clear symptoms of the disease and may have it in its late stage, a diagnosis known a priori and supported by our study.

In conclusion, we can say that the set of results presented in this paper attests that the QG method is an effective technique for the complex temporal pattern analysis like those found in EEGs from AD patients. It is worth mentioning that the subjects under study were not submitted to a definitive pathological diagnosis of AD as well as health controls. Some clinical features were not available at the time of this investigation. Moreover, the number of subjects is quite small limiting the extrapolations of the findings. Therefore, further research is necessary, with a larger and richer data set, to estimate the efficacy of the QG method in providing an early diagnostic of AD patients with only mild cognitive impairment.

Acknowledgments

The authors would like to thank Dr. Dennis Duke of Florida State University for providing the EEG database for the study.

Author Contributions

Conceptualization: Aruane M. Pineda, Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Data curation: Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Formal analysis: Aruane M. Pineda, Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Funding acquisition: Andriana S. L. O. Campanharo.

Investigation: Aruane M. Pineda, Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Methodology: Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Project administration: Aruane M. Pineda, Fernando M. Ramos, Andriana S. L. O. Campanharo.

Resources: Andriana S. L. O. Campanharo.

Software: Andriana S. L. O. Campanharo.

Supervision: Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Validation: Aruane M. Pineda, Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Visualization: Aruane M. Pineda, Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

Writing – original draft: Aruane M. Pineda, Fernando M. Ramos, Luiz Eduardo Betting, Andriana S. L. O. Campanharo.

References

1. Organization WH. Dementia: Switzerland: World Health Organization; 2012.
2. Budson A, Solomon P. Memory Loss, Alzheimer's Disease, and Dementia. New York: Elsevier; 2015.
3. Fjell AM, McEvoy L, Holland D, Dale AM, Walhovd KB, Initiative ADN, et al. What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in neurobiology*. 2014; 117:20–40. <https://doi.org/10.1016/j.pneurobio.2014.02.004> PMID: 24548606
4. Hyman BT, Van Hoesen GW, Damasio AR, Barnes CL. Alzheimer's disease: cell-specific pathology isolates the hippocampal formation. *Science*. 1984; 225(4667):1168–1170.
5. Feldman HH, Woodward M. The staging and assessment of moderate to severe Alzheimer disease. *Neurology*. 2005; 65. https://doi.org/10.1212/WNL.65.6_suppl_3.S10
6. Sharma N, Singh AN. Exploring biomarkers for Alzheimer's disease. *Journal of clinical and diagnostic research: JCDR*. 2016; 10(7):KE01. <https://doi.org/10.7860/JCDR/2016/18828.8166> PMID: 27630867
7. Karlawish J, Jack CR Jr, Rocca WA, Snyder HM, Carrillo MC. Alzheimer's disease: The next frontier—Special Report 2017. *Alzheimer's & Dementia*. 2017; 13(4):374–380. <https://doi.org/10.1016/j.jalz.2017.02.006>
8. Davatzikos C. Machine learning in neuroimaging: progress and challenges. *NeuroImage*. 2019; 197:652. <https://doi.org/10.1016/j.neuroimage.2018.10.003> PMID: 30296563
9. Yang S, Bornot JMS, Wong-Lin K, Prasad G. M/EEG-Based Bio-Markers to Predict the MCI and Alzheimer's Disease: A Review From the ML Perspective. *IEEE Transactions on Biomedical Engineering*. 2019; 66(10):2924–2935. <https://doi.org/10.1109/TBME.2019.2898871> PMID: 30762522
10. Horvath A, Szucs A, Csukly G, Sakovics A, Stefanics G, Kamondi A. EEG and ERP biomarkers of Alzheimer's disease: a critical review. *Frontiers in bioscience (Landmark edition)*. 2018; 23:183–220. <https://doi.org/10.2741/4587>
11. Acharya UR, Faust O, Kannathal N, Chua T, Laxminarayan S. Non-linear analysis of EEG signals at various sleep stages. *Computer Methods and Programs in Biomedicine*. 2005; 80. <https://doi.org/10.1016/j.cmpb.2005.06.011>
12. Alotaiby TN, Alshebeili SA, Alshawi T, Ahmad I, El-samie FEA. EEG seizure detection and prediction algorithms: a survey. *EURASIP Journal on Advances in Signal Processing*. 2014; 183.
13. Campanharo ASLO, Doescher E, Ramos FM. Automated EEG signals analysis using quantile graphs. In: *International Work-Conference on Artificial Neural Networks*. Springer; 2017. p. 95–103.
14. Campanharo ASLO, Doescher E, Ramos FM. Application of quantile graphs to the automated analysis of EEG signals. *Neural Processing Letters*. 2018; p. 1–16.
15. Tsolaki A, Kazis D, Kompatsiaris I, Kosmidou V, Tsolaki M. Electroencephalogram and Alzheimer's Disease: Clinical and Research Approaches. *International Journal of Alzheimer's Disease*. 2014; 2014. <https://doi.org/10.1155/2014/349249> PMID: 24868482
16. Locatelli T, Cursi M, Liberati D, Franceschi M, Comi G. EEG coherence in Alzheimer's disease. *Electroencephalography and clinical neurophysiology*. 1998; 106(3):229–237. [https://doi.org/10.1016/s0013-4694\(97\)00129-6](https://doi.org/10.1016/s0013-4694(97)00129-6) PMID: 9743281
17. Korner TW. Fourier analysis. Cambridge: Cambridge University Press; 1988.

18. Fiscon G, Weitschek E, Felici G, Bertolazzi P, De Salvo S, Bramanti P, et al. Alzheimer's disease patients classification through EEG signals processing. In: 2014 IEEE Symposium on Computational Intelligence and Data Mining (CIDM). IEEE; 2014. p. 105–112.
19. Smits FM, Porcaro C, Cottone C, Cancelli A, Rossini PM, Tecchio F. Electroencephalographic fractal dimension in healthy ageing and Alzheimer's disease. *PLoS one*. 2016; 11(2). <https://doi.org/10.1371/journal.pone.0149587>
20. Percival DB, Walden A. Wavelet methods for time series analysis. Cambridge: Cambridge University Press; 2000.
21. Yagneswaran S, Baker M, Petrosian A. Power frequency and wavelet characteristics in differentiating between normal and Alzheimer EEG. In: Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society][Engineering in Medicine and Biology. vol. 1. IEEE; 2002. p. 46–47.
22. Ghorbanian P, Devilbiss DM, Hess T, Bernstein A, Simon AJ, Ashrafiun H. Exploration of EEG features of Alzheimer's disease using continuous wavelet transform. *Medical & biological engineering & computing*. 2015; 53(9):843–855. <https://doi.org/10.1007/s11517-015-1298-3>
23. Fiscon G, Weitschek E, Cialini A, Felici G, Bertolazzi P, De Salvo S, et al. Combining EEG signal processing with supervised methods for Alzheimer's patients classification. *BMC medical informatics and decision making*. 2018; 18(1):35. <https://doi.org/10.1186/s12911-018-0613-y> PMID: 29855305
24. Kantz H, Schreiber T. Nonlinear time series analysis. Cambridge: Cambridge University Press; 2003.
25. Stam C, Jelles B, Achtereekte H, Van Birgelen J, Slaets J. Diagnostic usefulness of linear and nonlinear quantitative EEG analysis in Alzheimer's disease. *Clinical Electroencephalography*. 1996; 27(2). <https://doi.org/10.1177/155005949602700205> PMID: 8681465
26. Kumar SS, Kasthuri N. EEG seizure classification based on exploiting phase space reconstruction and extreme learning. *Cluster Computing*. 2017; p. 1–11.
27. Ubeyli ED. Analysis of EEG signals by combining eigenvector methods and multiclass support vector machines. *Computers in Biology and Medicine*. 2011; 38.
28. Ubeyli ED, Guler I. Features extracted by eigenvector methods for detecting variability of EEG signals. *Computers in Biology and Medicine*. 2007; 28.
29. Ridouh A, Boutana D, Bourenane S. EEG Signals Classification Based on Time Frequency Analysis. *Journal of Circuits, Systems and Computers*. 2017; 26. <https://doi.org/10.1142/S0218126617501985>
30. Zhang Y, Liu B, Ji X, Huang D. Classification of EEG Signals Based on Autoregressive Model and Wavelet Packet Decomposition. *Neural Processing Letters*. 2016; 45.
31. Campanharo ASLO, Sirel MI, Malmgren RD, Ramos FM, Amaral LAN. Duality between Time Series and Networks. *PLoS ONE*. 2011; 6. <https://doi.org/10.1371/journal.pone.0023378> PMID: 21858093
32. Campanharo ASLO, Ramos FM. Hurst exponent estimation of self-affine time series using quantile graphs. *Physica A*. 2016; 444. <https://doi.org/10.1016/j.physa.2015.09.094>
33. AM, Ramos FM, Betting LE, Campanharo AS. Use of Complex Networks for the Automatic Detection and the Diagnosis of Alzheimer's Disease. In: International Work-Conference on Artificial Neural Networks. Springer; 2019. p. 115–126.
34. Morris AS, Langari R. Measurement and Instrumentation. San Diego, CA: Academic Press; 2012.
35. Pineda AM, Campanharo ASLO, Ramos FM, Betting LE. Code from: Quantile Graphs for EEG-Based Diagnosis of Alzheimer's Disease; 2020. <https://osf.io/6w5fk/>.
36. Sporns O. The human connectome: a complex network. *Annals of the New York Academy of Sciences*. 2011; 1224(1):109–125. <https://doi.org/10.1111/j.1749-6632.2010.05888.x> PMID: 21251014
37. Bounova G. Matlab Tools for Network Analysis octave-networks-toolbox; 2011. <http://strategic.mit.edu/>.
38. Saramäki J, Kivela M, Onnela JP, Kaski K, Kertesz J. Generalizations of the clustering coefficient to weighted complex networks. *Physical Review E*. 2007; 75(2):027105. <https://doi.org/10.1103/PhysRevE.75.027105>
39. Freeman LC. A set of measures of centrality based on betweenness. *Sociometry*. 1977; 40. <https://doi.org/10.2307/3033543>
40. Costa LF, Rodrigues FA, Travieso G, Villas PR. Characterization of complex networks. *Advances in Physics*. 2007; 56. <https://doi.org/10.1080/00018730601170527>
41. Newman MEJ. Modularity and community structure in networks. *Proc Natl Acad Sci USA*. 2006; 103. <https://doi.org/10.1073/pnas.0601602103>
42. Boccaletti S, Ivanchenko M, Latora V, Pluchino A, Rapisarda A. Detecting complex network modularity by dynamical clustering. *Physical Review E*. 2007; 75(4):045102. <https://doi.org/10.1103/PhysRevE.75.045102>

43. Guimera R, Amaral LAN. Cartography of complex networks: modules and universal roles. *Journal of Statistical Mechanics: Theory and Experiment*. 2005; 2005(02):P02001. <https://doi.org/10.1088/1742-5468/2005/02/P02001>
44. Sales-Pardo M, Guimerà R, Amaral LAN. Extracting the hierarchical organization of complex systems. *Proc Natl Acad Sci USA*. 2007; 104. <https://doi.org/10.1073/pnas.0703740104> PMID: 17881571
45. Cvetkovic DM, Doob M, Sachs H, et al. *Spectra of graphs*. vol. 10. Academic Press, New York; 1980.
46. Fath-Tabar G, Ashrafi A, Gutman I. Note on Estrada and *L*-Estrada indices of graphs. *Bulletin, Classe des Sciences Mathématiques et Naturelles, Sciences mathématiques*. 2009; 139(34):1–16.
47. Pritchard WS, Duke DW, Coburn KL, Moore NC, Tucker KA, Jann MW, et al. EEG-based, neural-net predictive classification of Alzheimer's disease versus control subjects is augmented by non-linear EEG measures. *Electroencephalography and clinical Neurophysiology*. 1994; 91(2):118–130. [https://doi.org/10.1016/0013-4694\(94\)90033-7](https://doi.org/10.1016/0013-4694(94)90033-7) PMID: 7519141
48. Pritchard WS, Duke DW, Coburn KL. Altered EEG dynamical responsivity associated with normal aging and probable Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders*. 1991; 2(2):102–105. <https://doi.org/10.1159/000107183>
49. Campanharo ASLO, Ramos FM, Pineda AM, Betting LE. Data from: Quantile Graphs for EEG-Based Diagnosis of Alzheimer's Disease; 2020. <https://osf.io/s74qf/>.
50. Hajian-Tilaki K. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian Journal of Internal Medicine*. 2013; 4. PMID: 24009950
51. Obuchowski NA, Bullen JA. Receiver operating characteristic (ROC) curves: review of methods with applications in diagnostic medicine. *Physics in Medicine & Biology*. 2018; 63.
52. Lee J, Kim Y, Jeong Y, Na DL, Kim JW, Lee KH, et al. Inference of brain pathway activities for Alzheimer's disease classification. *BMC medical informatics and decision making*. 2015; 15(1):S1. <https://doi.org/10.1186/1472-6947-15-S1-S1> PMID: 26044913
53. Jeong J. EEG dynamics in patients with Alzheimer's disease. *Clinical neurophysiology*. 2004; 115(7):1490–1505. <https://doi.org/10.1016/j.clinph.2004.01.001> PMID: 15203050
54. Al-Nuaimi AHH, Jammeh E, Sun L, Ifeachor E. Complexity measures for quantifying changes in electroencephalogram in Alzheimer's disease. *Complexity*. 2018; 2018. <https://doi.org/10.1155/2018/8915079>
55. Bennys K, Rondouin G, Vergnes C, Touchon J. Diagnostic value of quantitative EEG in Alzheimer's disease. *Neurophysiologie Clinique/Clinical Neurophysiology*. 2001; 31(3):153–160. [https://doi.org/10.1016/s0987-7053\(01\)00254-4](https://doi.org/10.1016/s0987-7053(01)00254-4) PMID: 11488226
56. Helkala EL, Hänninen T, Hallikainen M, Könönen M, Laakso M, Hartikainen P, et al. Slow-wave activity in the spectral analysis of the electroencephalogram and volumes of hippocampus in subgroups of Alzheimer's disease patients. *Behavioral neuroscience*. 1996; 110(6):1235. <https://doi.org/10.1037//0735-7044.110.6.1235> PMID: 8986328
57. Henry JC. *Electroencephalography: basic principles, clinical applications, and related fields*. *Neurology*. 2006; 67(11):2092–2092. <https://doi.org/10.1212/01.wnl.0000243257.85592.9a>
58. Herrmann CS, Grigutsch M, Busch NA. 11 EEG oscillations and wavelet analysis. *Event-related potentials: A methods handbook*. 2005; p. 229.
59. MATLAB version 9.3.0.713579 (R2017b); 2017.
60. Aisen PS, Cummings J, Jack CR, Morris JC, Sperling R, Frölich L, et al. On the path to 2025: understanding the Alzheimer's disease continuum. *Alzheimer's research & therapy*. 2017; 9(1):60. <https://doi.org/10.1186/s13195-017-0283-5>
61. Lopez-de Ipina K, Martinez-de Lizarduy U, Calvo PM, Mekyska J, Beitia B, Barroso N, et al. Advances on automatic speech analysis for early detection of Alzheimer disease: a non-linear multi-task approach. *Current Alzheimer Research*. 2018; 15(2):139–148. <https://doi.org/10.2174/1567205014666171120143800> PMID: 29165084
62. Loewenstein DA, Barker WW, Chang JY, Apicella A, Yoshii F, Kothari P, et al. Predominant left hemisphere metabolic dysfunction in dementia. *Archives of Neurology*. 1989; 46(2):146–152. <https://doi.org/10.1001/archneur.1989.00520380046012> PMID: 2783845