

**USO DE REDES COMPLEXAS NO ESTUDO E NO DIAGNÓSTICO
DA DOENÇA DE ALZHEIMER**

Aruane Mello Pineda

Dissertação apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de mestre em Biometria.

BOTUCATU
São Paulo - Brasil
Julho – 2019

**USO DE REDES COMPLEXAS NO ESTUDO E NO DIAGNÓSTICO
DA DOENÇA DE ALZHEIMER**

Aruane Mello Pineda

Orientadora: Prof^ª. Dr^ª. **Andriana S. L. O. Campanharo**

Dissertação apresentada à Universidade Estadual Paulista “Júlio de Mesquita Filho” para obtenção do título de mestre em Biometria.

BOTUCATU
São Paulo - Brasil
Julho – 2019

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉC. AQUIS. TRATAMENTO DA INFORM.
DIVISÃO TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - CÂMPUS DE BOTUCATU - UNESP
BIBLIOTECÁRIA RESPONSÁVEL: LUCIANA PIZZANI-CRB 8/6772

Pineda, Aruane Mello.

Uso de redes complexas no estudo e no diagnóstico da
doença de Alzheimer / Aruane Mello Pineda. - Botucatu,
2019

Dissertação (mestrado) - Universidade Estadual Paulista
"Júlio de Mesquita Filho", Instituto de Biociências de
Botucatu

Orientador: Andriana Susana Lopes de Oliveira
Campanharo

Capes: 33004064

1. Alzheimer, Doença de. 2. Análise de séries
temporais. 3. Eletroencefalografia.

Palavras-chave: Doença de Alzheimer; EEG; Mapeamento de
séries temporais em redes complexas; Redes complexas.

ATA DA DEFESA PÚBLICA DA DISSERTAÇÃO DE MESTRADO DE ARUANE MELLO PINEDA, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA, DO INSTITUTO DE BIOCÊNCIAS - CÂMPUS DE BOTUCATU.

Aos 16 dias do mês de julho do ano de 2019, às 09:00 horas, no(a) Departamento de Bioestatística, reuniu-se a Comissão Examinadora da Defesa Pública, composta pelos seguintes membros: Profa. Dra. ANDRIANA SUSANA LOPES DE OLIVEIRA CAMPANHARO - Orientador(a) do(a) Departamento de Bioestatística / Instituto de Biociências de Botucatu - UNESP, Prof. Dr. FERNANDO MANUEL RAMOS do(a) Departamento de Computação Aplicada / INPE, Prof. Dr. LUIZ EDUARDO GOMES GARCIA BETTING do(a) Departamento de Neurologia, Psicologia e Psiquiatria / Faculdade de Medicina de Botucatu - UNESP, sob a presidência do primeiro, a fim de proceder a arguição pública da DISSERTAÇÃO DE MESTRADO de ARUANE MELLO PINEDA, intitulada **Uso de redes complexas no estudo e no diagnóstico da doença de Alzheimer**. Após a exposição, a discente foi arguida oralmente pelos membros da Comissão Examinadora, tendo recebido o conceito final: _____
Aprovada. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelos membros da Comissão Examinadora.

Profa. Dra. ANDRIANA SUSANA LOPES DE OLIVEIRA CAMPANHARO

Prof. Dr. FERNANDO MANUEL RAMOS

Prof. Dr. LUIZ EDUARDO GOMES GARCIA BETTING

Dedico este trabalho a Deus, a minha família, meu namorado e a minha orientadora que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida.

Agradecimentos

Agradeço a todos que contribuíram para a conclusão deste trabalho e realização deste sonho, especialmente:

Aos meus pais, Dora e Antonio, a minha irmã Fernanda e meu namorado Claudio.

A minha orientadora, Dr^aAndriana Susana Lopes de Oliveira Campanharo, pela oportunidade de pesquisa, orientação, paciência e compreensão durante esses anos.

Aos meus colegas do departamento Eduardo Ribeiro, Felipe Teles e Gustavo Tomanik pela paciência e cumplicidade.

À banca, Prof. Dr. Luiz Eduardo Gomes Garcia Betting e Prof. Dr. Fernando Manuel Ramos pelas sugestões para o aprimoramento deste trabalho.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Sumário

| | Página |
|--|-------------|
| LISTA DE FIGURAS | vii |
| LISTA DE TABELAS | xi |
| RESUMO | xiii |
| SUMMARY | xv |
| 1 INTRODUÇÃO | 1 |
| 1.1 A doença de Alzheimer | 1 |
| 1.2 EletroEncefaloGrafia (EEG) | 3 |
| 1.3 Análise de Séries Temporais | 4 |
| 1.4 Novos métodos de análise de séries temporais | 5 |
| 1.5 Estrutura do texto | 6 |
| 2 REDES COMPLEXAS | 8 |
| 2.1 Aspectos históricos | 10 |
| 2.2 Propriedades topológicas | 11 |
| 2.2.1 Grau de conectividade | 11 |
| 2.2.2 Força de conectividade | 12 |
| 2.2.3 Coeficiente de agrupamento médio (C) | 13 |
| 2.2.4 Modularidade (M) | 14 |
| 2.2.5 Comprimento do salto médio (Δ) | 16 |
| 2.2.6 Grau de Intermediação (B) | 17 |

| | |
|--|-----------|
| | vi |
| 3 METODOLOGIA | 19 |
| 3.1 Mapeamento de séries temporais em Redes Complexas | 19 |
| 3.2 Análise de Alzheimer via dados de EEG | 21 |
| 4 RESULTADOS | 23 |
| 4.1 Distinção entre envelhecimento e a DA | 23 |
| 4.2 Desempenho dos caracterizadores com base nos canais de EEG | 36 |
| 4.2.1 Investigação dos possíveis estágios da DA | 39 |
| 4.3 Discussão | 44 |
| 5 CONCLUSÕES | 46 |
| REFERÊNCIAS BIBLIOGRÁFICAS | 47 |

Lista de Figuras

| | Página |
|--|--------|
| 1 Ilustração da disposição dos canais de EEG. Fonte: https://www.physionet.org/pn6/chbmit/ | 3 |
| 2 Exemplo de rede direcionada e com peso, e sua matriz de adjacência com peso W | 9 |
| 3 Representação esquemática das relações entre os tipos de redes. Fonte: Adaptada de (Rodrigues, 2007). | 9 |
| 4 O problema das sete pontes de Königsberg. Representação de Euler por meio de um grafo para o problema. Fonte: (Rodrigues, 2007). | 11 |
| 5 Exemplo de rede com peso e com direção, constituída de 4 vértices. Com base nas Equações 11 e 12, o coeficiente de agrupamento médio dessa rede é dado por $C \approx 0,31$ | 14 |
| 6 Exemplo de rede modular sem peso e sem direção, constituída de 6 vértices, 7 arestas e 2 módulos. Com base na Equação 13, a modularidade dessa rede é dada por $\mathcal{M}_p = 0,35$ | 15 |
| 7 Exemplo de rede com peso e com direção, constituída de 3 vértices. Com base na Equação 15, o comprimento do salto médio dessa rede é dado por $\Delta \approx 0,94$ | 16 |
| 8 Exemplo de rede sem peso e sem direção, constituída de 5 vértices e 5 arestas. O grau de intermediação para os vértices 1, 2, 3, 4 e 5 são dados por, 0,08, 0,16, 0,16, 0,58 e 0,00, respectivamente. Tal rede possui grau de intermediação médio dado por 0,58. | 18 |

- 9 Ilustração do mapeamento proposto por Campanharo *et al.* (Campanharo *et al.*, 2011; Campanharo & Ramos, 2016a) para uma série temporal com $T = 15$ pontos, $Q = 5$ quantis e $k = 1, 2$ e 3 , mapeada em três redes com cinco vértices cada. Observe que, independente do valor de k escolhido, tal mapeamento produz redes com peso, direção e conexas. Os intervalos dos quantis são dados por $[x(0), x(3)[$, $[x(3), x(6)[$, $[x(6), x(9)[$, $[x(9), x(12)[$ e $[x(12), x(15)]$ para os dados ordenados, isto é, $[-8, 740, -4, 912[$, $[-4, 912, 0, 128[$, $[0, 128, 2, 743[$, $[2, 743, 5, 000[$ e $[5, 000, 6, 571]$. Fonte: (Pineda *et al.*, 2019). 20
- 10 Exemplos de séries de EEG para cada um dos conjuntos em estudo (canal F_7 - lobo frontal). De cima para baixo: A (paciente saudável, olhos abertos), B (paciente saudável, olhos fechados), C (paciente doente, olhos abertos) e D (paciente doente, olhos fechados). 26
- 11 $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , (canal F_7 - lobo frontal) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos), B (pacientes saudáveis, olhos fechados), C (pacientes doentes, olhos abertos) e D (pacientes doentes, olhos fechados). 26
- 12 Boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, (canal F_7 - lobo frontal) para os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados) e $k = 9, 6, 10$ e 6 , respectivamente. Pacientes em diferentes condições de saúde possuem medianas distintas (linhas centrais de cada caixa), dadas por $0,2077, 0,2033, 0,2248$ e $0,2232$ para $C(k)$; $0,0605, 0,0728, 0,0145$ e $0,0072$ para $M(k)$; $4,8010, 4,2190, 6,6790$ e $6,6510$ para $\Delta(k)$ e $0,0111, 0,0129, 0,0061$ e $0,0054$ para $B(k)$, respectivamente. 27
- 13 Exemplos de séries de EEG para cada um dos conjuntos em estudo (canal F_8 - lobo frontal). De cima para baixo: A (paciente saudável, olhos abertos), B (paciente saudável, olhos fechados), C (paciente doente, olhos abertos), D (paciente doente, olhos fechados). 29

- 14 $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , (canal F_8 - lobo frontal) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos), B (pacientes saudáveis, olhos fechados), C (pacientes doentes, olhos abertos) e D (pacientes doentes, olhos fechados). 29
- 15 Boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, (canal F_8 - lobo frontal) para os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados) e $k = 7, 5, 8$ e 5 , respectivamente. Pacientes em diferentes condições de saúde possuem medianas distintas (linhas centrais de cada caixa), dadas por 0,2102, 0,2055, 0,2221 e 0,2213 para $C(k)$; 0,0398, 0,0619, 0,0080 e 0,0108 para $M(k)$; 4,9010, 4,3730, 6,3800 e 6,4520 para $\Delta(k)$ e 0,0093, 0,0113, 0,0058 e 0,0061 para $B(k)$, respectivamente. 30
- 16 Exemplos de séries de EEG para cada um dos conjuntos em estudo (canal P_z - lobo parietal). De cima para baixo: A (paciente saudável, olhos abertos), B (paciente saudável, olhos fechados), C (paciente doente, olhos abertos), D (paciente doente, olhos fechados). 32
- 17 $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , (canal P_z - lobo parietal) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos), B (pacientes saudáveis, olhos fechados), C (pacientes doentes, olhos abertos) e D (pacientes doentes, olhos fechados). 32
- 18 Boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, (canal P_z - lobo parietal) para os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados) e $k = 12, 13, 11$ e 12 , respectivamente. Pacientes em diferentes condições de saúde possuem medianas distintas (linhas centrais de cada caixa), dadas por 0,2133, 0,2096, 0,2225 e 0,2241 para $C(k)$; 0,0245, 0,0397, 0,0000 e 0,0000 para $M(k)$; 4,8000, 4,7260, 6,8750 e 6,8300 para $\Delta(k)$ e 0,0074, 0,0084, 0,0046 e 0,0043 para $B(k)$, respectivamente. 33

| | | |
|----|--|----|
| 19 | Matrizes de transição de Markov para $k = 11$ (canal P_z - lobo parietal) dos conjuntos (a) (saudáveis, olhos abertos), (b) (saudáveis, olhos fechados), (c) (doentes, olhos abertos) e (d) (doentes, olhos fechados). | 35 |
| 20 | Desempenho dos caracterizadores C , M , Δ e B com base no valores de \hat{A}_{ROC} para os 19 canais de EEG, na distinção entre pacientes sadios e doentes. Com base nas combinações entre os conjuntos A e C, B e D, A e D e B e C, foi calculado a média dos valores de \hat{A}_{ROC} que varia de 0,7143 a 1,0000. | 37 |
| 21 | Desempenho dos caracterizadores C , M , Δ e B com base nos valores de A_{ROC} para os 19 canais de EEG, na distinção entre pacientes sadios e doentes com olhos abertos. Com base nas combinações entre os conjuntos A e C, foi calculado os valores de A_{ROC} que varia de 0,6736 a 1,0000. . . | 38 |
| 22 | Desempenho dos caracterizadores C , M , Δ e B com base nos valores de A_{ROC} para os 19 canais de EEG, na distinção entre pacientes sadios e doentes com olhos fechados. Com base nas combinações entre os conjuntos B e D, foi calculado os valores de A_{ROC} que varia de 0,6667 a 1,0000. . . | 39 |
| 23 | $C(\theta)$, $C(\beta)$, $C(\alpha)$ e $C(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos). | 41 |
| 24 | $M(\theta)$, $M(\beta)$, $M(\alpha)$ e $M(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos). | 42 |
| 25 | $\Delta(\theta)$, $\Delta(\beta)$, $\Delta(\alpha)$ e $\Delta(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos). | 43 |
| 26 | $B(\theta)$, $B(\beta)$, $B(\alpha)$ e $B(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos). | 44 |

Lista de Tabelas

| | | Página |
|---|--|--------|
| 1 | Valores de t_i , k_i^{tot} , k_i^{\leftrightarrow} e C_i para cada um dos vértices da rede da Figura 5. | 14 |
| 2 | Cálculo do grau de intermediação para o vértice 4. | 18 |
| 3 | Valores de k escolhidos para cada canal e cada classificador de rede em estudo. | 24 |
| 4 | Comparação estatística entre as médias de C , M , Δ e $B(k)$ para $k = 9$, 6 e 10 e 6, (canal F_7 - lobo frontal) respectivamente entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados), com base no teste estatístico ANOVA. | 28 |
| 5 | Áreas sob as curvas ROC de C , M , Δ e B para $k = 9$, 6, 10 e 6 (canal F_7 - lobo frontal) entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados). | 28 |
| 6 | Comparação estatística entre as médias de C , M , Δ e B para $k = 7$, 5, 8 e 5, (canal F_8 - lobo frontal) respectivamente entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados), com base no teste estatístico ANOVA. | 31 |
| 7 | Áreas sob as curvas ROC de C , M , Δ e B para $k = 7$, 5, 8 e 5 (canal F_8 - lobo frontal) entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados). | 31 |

| | | |
|---|--|----|
| 8 | Comparação estatística entre as médias de C , M , Δ e B para $k = 12$, 13, 11 e 12, (canal P_z - lobo parietal) respectivamente entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados), com base no teste estatístico ANOVA. | 34 |
| 9 | Áreas sob as curvas ROC de C , M , Δ e B para $k = 12$, 13, 11 e 12 (canal P_z - lobo parietal) entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados). | 34 |

USO DE REDES COMPLEXAS NO ESTUDO E NO DIAGNÓSTICO DA DOENÇA DE ALZHEIMER

Autora: ARUANE MELLO PINEDA

Orientadora: Prof^ª. Dr^ª. ANDRIANA S. L. O. CAMPANHARO

RESUMO

O Alzheimer é uma doença degenerativa do cérebro, incurável, que se agrava ao longo do tempo e atinge sobretudo pessoas entre 65 e 90 anos. A doença de Alzheimer (DA) é a principal demência entre os idosos e caracteriza-se por perda de funções cognitivas (memória, orientação, comportamento e linguagem), causada pela morte de células cerebrais. Atualmente, o diagnóstico definitivo da DA é feito por meio de um exame do tecido cerebral obtido por biopsia ou necropsia. Como somente após a morte do paciente pode-se ter a certeza que o mesmo tinha a DA, seu diagnóstico é feito utilizando exames, avaliações e excluindo-se outras causas de demência do seu histórico clínico. Em paralelo, estudos têm sido desenvolvidos para o estudo da DA com base de dados de EletroEncefaloGramma (EEG), e nesse sentido, diversos métodos de análise de dados de EEG têm sido propostos. Contudo, o estudo da DA por meio de dados de EEG é ainda um desafio, e conseqüentemente, faz-se necessária a preposição de novos métodos com o intuito de capturar informações

adicionais da doença. Nesse sentido, utilizamos o mapeamento de uma série temporal em uma rede complexa no estudo da dinâmica de séries temporais de EEG de pacientes com a DA. Mais especificamente, na distinção entre envelhecimento e a DA e na identificação, dos estágios da DA em pacientes doentes a partir, das ondas cerebrais, alfa, beta, teta e delta.

Palavras-chave: EEG, Redes Complexas, Doença de Alzheimer, Mapeamento de Séries Temporais em Redes Complexas.

USE OF COMPLEX NETWORKS IN THE STUDY AND DIAGNOSIS OF ALZHEIMER'S DISEASE

Author: ARUANE MELLO PINEDA

Adviser: Prof^a. Dr^a. ANDRIANA S. L. O. CAMPANHARO

SUMMARY

Alzheimer's is a degenerative brain disease, incurable, which aggravates over time and mainly affects people between 65 and 90 years. Alzheimer's disease (AD) is the leading dementia among the elderly and is characterized by cognitive functions loss (memory, orientation, behavior and language) caused by the death of brain cells. Currently, confirmatory diagnosis of AD can only be made through the examination the brain tissue obtained by biopsy or necropsy. Considering that only after the patient's death it is possible to be sure that he or she had AD, the approximate diagnosis is made excluding other causes of dementia. In parallel, studies have been developed for the study of AD with the use of Electroencephalography (EEG), and in this sense, several methods of EEG data analysis have been proposed. However, the study of AD with the use of EEG data is still a challenge, and consequently, it is necessary the proposal of new methods to capture additional information about

the disease. In this sense, we used the map from a time series to a network, recently proposed by Campanharo, in a novel application, that is, in the study of EEG time series of patients with AD. More specifically, to distinguish aging and AD and to identify the stages of AD in unhealthy patients based on alpha, beta, theta and delta brain waves information.

Keyword: EEG, Complex Networks, Alzheimer's Disease, Map from a Time Series to a Network.

1 INTRODUÇÃO

1.1 A doença de Alzheimer

A doença de Alzheimer (DA) pode ser entendida como uma demência degenerativa e progressiva do Sistema Nervoso Central, irreversível e de causa desconhecida. Tal doença é caracterizada, principalmente, pela deterioração intelectual acelerada e por perda de memória e desorientação no tempo e no espaço (Budson & Solomon, 2015). A DA é uma demência que incide em idosos com mais de 65 anos e afeta, aproximadamente, 25 milhões de indivíduos em todo o mundo (Org, 2012). Nos EUA, a prevalência da DA foi estimada em 4,5 milhões em 2.000, e projetada para 13,2 milhões em 2.050 (Husain & Garrett, 2005). No Brasil, o número de idosos com a doença é de aproximadamente 1,2 milhões e, a cada ano, surgem em média 100 mil novos casos (Gutierrez et al., 2014). A DA constitui aproximadamente 70% de todos os casos de demência, sendo que a sua incidência aumenta com a idade, dobrando a cada cinco/dez anos. Para as pessoas entre as idades de 65-69, 70-74, 75-79, 80-84 e 85, a incidência da DA tem sido estimada em 0,6%; 1,0%; 2,0%; 3,3% e 8,4%, respectivamente. Sua prevalência também aumenta exponencialmente com a idade, passando de 3% entre os 65-74 anos, para quase 50% entre os mais de 85 anos (Budson & Solomon, 2015).

O desenvolvimento da DA é dividido em três fases, a saber, inicial, intermediário (leve e grave) e terminal (Feldman & Woodward, 2005). Na fase inicial, ocorrem as primeiras mudanças típicas da doença. Os neurônios do lobo temporal e do hipocampo são afetados. Registram-se episódios de esquecimento que não atrapalham a convivência e o doente é ainda independente. Na fase intermediária leve,

outras zonas do córtex cerebral são afetadas e são registrados episódios de perda de memória recente, dificuldades para aprender e reter novas informações, distúrbios de linguagem, dificuldade progressiva na realização de tarefas diárias, desorientação e apatia. Na fase intermediária grave, os neurônios dos lobos frontal e parietal são afetados e é observada a dificuldade de socialização e perda de memória mais intensa. O paciente torna-se completamente incapaz de aprender e reter novas informações e iniciam-se as dificuldades de locomoção. A fase terminal é caracterizada pela atrofia geral do cérebro, onde observa-se a presença de emaranhados neurofibrilares, a diminuição da densidade sináptica e a morte de neurônios. Nessa fase, o paciente encontra-se completamente dependente de cuidados evoluindo gradativamente para o estado vegetativo. A DA não tem cura já que os neurônios perdidos não podem ser recuperados. Contudo, o diagnóstico e o tratamento precoce podem retardar sua evolução, tratar seus sintomas e as alterações de comportamento em pacientes com essa doença (Feldman & Woodward, 2005).

Atualmente, o diagnóstico definitivo da DA pode ser feito apenas por meio de um exame do tecido cerebral obtido por biopsia ou necropsia. Como somente após a morte do paciente pode-se ter a certeza que o mesmo tinha a DA, seu diagnóstico é feito excluindo-se outras causas de demência pelo histórico clínico do paciente, tais como, doenças preexistentes, cirurgias, uso de medicações, depressão e perda de memória associada à idade. Métodos como Imagem por Ressonância Magnética (IRM), Tomografia Computadorizada e Tomografia por Emissão de Pósitrons (PET) têm sido utilizadas para auxiliar profissionais no processo de diagnóstico (Durongbhan et al., 2019). Em paralelo, estudos têm sido desenvolvidos para a investigação da DA com base em informações da atividade cerebral. Mais especificamente, por meio de dados obtidos pela técnica da EletroEncefalografia (EEG).

1.2 Eletroencefalografia (EEG)

A Eletroencefalografia (EEG) consiste no registro gráfico das correntes elétricas desenvolvidas no encéfalo, obtido por meio de eletrodos aplicados no couro cabeludo, na superfície encefálica, ou até mesmo dentro da substância encefálica, que permite por meio de um computador o registro das atividades dos neurônios. Cada eletrodo é posicionado em áreas do cérebro, denominadas lobos cerebrais, cada qual com funções diferenciadas e especializadas, a saber, lobo frontal, parietal, temporal e occipital. A técnica de EEG é comumente utilizada no diagnóstico de doenças, tais como, distúrbios do sono, esquizofrenia, epilepsia e a DA (Acharya et al., 2005; Santos-Mayo et al., 2016; Tsolaki et al., 2014). Na base de dados utilizada neste trabalho foram fornecidas informações sobre 19 canais de EEG e a nomenclatura dos canais é dada de acordo com a região que estão localizados, Fp_1 , Fp_2 , F_7 , F_3 , F_z , F_4 e F_8 (lobo frontal), T_3 , T_4 , T_5 e T_6 (lobo temporal), C_3 , C_z , C_4 (região central), P_3 , P_z e P_4 (lobo parietal) e O_1 e O_2 (lobo occipital) (Fig. 1).

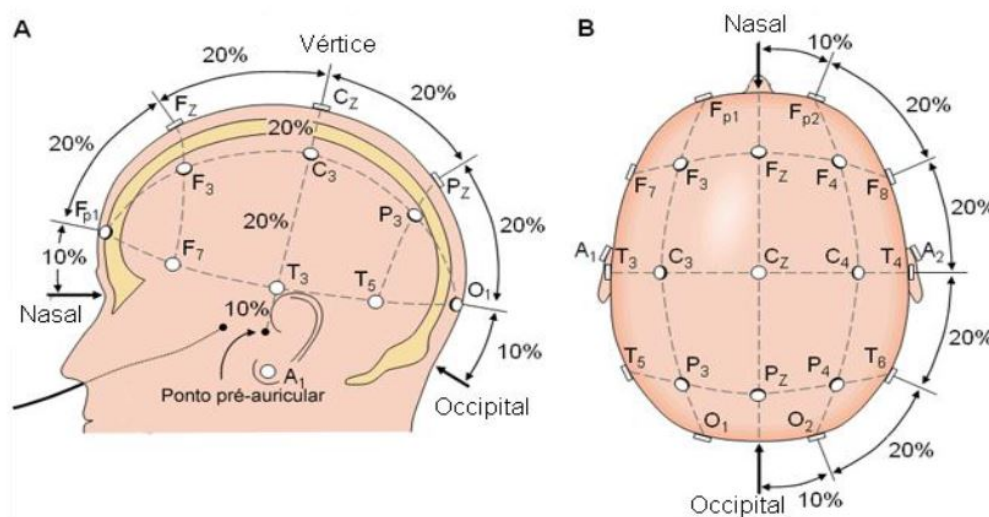


Figura 1: Ilustração da disposição dos canais de EEG.

Fonte: <https://www.physionet.org/pn6/chbmit/>

A análise visual dos dados de EEG constitui uma atividade exaustiva, que demanda tempo e especialistas. Além disso, o diagnóstico de doenças por meio de

tal análise é subjetiva, já que as opiniões desses especialistas não são necessariamente as mesmas. O EEG apesar de baixa resolução espacial apresenta como grandes vantagens o baixo custo, elevada resolução temporal e ampla disponibilidade. Desta forma, a análise quantitativa deste exame pode acrescentar informações importantes no diagnóstico de doenças em indivíduos. Portanto, a avaliação automática de tal atividade é de grande importância.

1.3 Análise de Séries Temporais

Em estatística e em processamento de sinais, uma série temporal é definida como uma sequência de medidas, obtida em tempos sucessivos e espaçada em intervalos de tempo uniformes. A caracterização da dinâmica de um sistema a partir de séries temporais experimentais e sintéticas é um problema de interesse contínuo em uma grande variedade de campos de investigação científica. A análise de séries temporais é usada, por exemplo, na classificação e identificação de condições patológicas de pacientes com base em dados de eletroencefalograma, respiração, batimento cardíaco e pressão arterial pulmonar, o que permite inúmeras possibilidades de procedimentos terapêuticos não-invasivos e o aumento da capacidade de diagnóstico (Campanharo, 2011; Amaral et al., 2004, 2001; Lo et al., 2002; Ivanov et al., 2000, 1998).

Atualmente, existe um grande número de técnicas de análise de séries temporais que varia desde métodos tempo-frequência, tais como as transformadas de Fourier e Wavelet (Yagneswaran et al., 2002; Korner, 1988; Percival & Walden, 2000), redes neurais artificiais (Pritchard et al., 1994), à métodos não-lineares, tais como, a reconstrução do espaço de fases, dimensão de correlação, o expoente de Lyapunov, as entropias e a complexidade de Lempel-Ziv (Stam et al., 1996; Strogatz, 1994; Kantz & Schreiber, 2003; Houmani et al., 2018; Adeli et al., 2005). Tais técnicas permitem pesquisadores sumarizar as características de uma série temporal, e desta forma, determinar a dinâmica subjacente de um sistema ou prever como o mesmo evolui com o tempo. Contudo, tais técnicas são amplamente utilizadas na análise de séries

temporais estacionárias, e neste caso, produzem resultados satisfatórios. Do ponto de vista estatístico, para produzirem resultados confiáveis, as mesmas fazem uso de um número elevado de amostras (realizações) de um mesmo tipo de série, o que inviabiliza a análise de muitas séries temporais experimentais, já que em geral, tais séries são não-estacionárias e obtidas diretamente de experimentos com um número reduzido de amostras (Campanharo, 2011; Eckmann & Ruelle, 1992).

Mais especificamente, a transformada de Fourier é a mais rápida quando comparada com os outros métodos disponíveis. No entanto, a mesma não permite a análise de sinais não-estacionários. Além disso, essa técnica não pode ser empregada na análise de sinais curtos e possui alta sensibilidade a ruídos (Al-Fahoum & Al-Fraihat, 2014). A transformada Wavelet foi apresentada como uma solução na análise de sinais não-estacionários. No entanto, a performance dessa técnica está condicionada à escolha adequada da ondeleta mãe na decomposição do sinal em diferentes escalas de frequência e tempo (Campanharo et al., 2018). As redes neurais artificiais são capazes de identificar padrões em séries temporais e dividi-las em classes segundo um critério. No entanto, para garantir que os resultados sejam estatisticamente precisos faz se necessário o uso de uma quantidade grande de amostras para serem testadas e treinadas (Cannady, 1998; Tu, 1996). Métodos não-lineares permitem a análise de sinais longos e estacionários, porém, são métodos altamente sensíveis à presença de ruídos e à escolha de diversos parâmetros, tais como, o atraso temporal e a dimensão de imersão (Strogatz, 1994).

1.4 Novos métodos de análise de séries temporais

A caracterização da dinâmica de séries temporais é um desafio constante, e conseqüentemente, novos métodos têm sido propostos no meio científico com o intuito de capturar informações adicionais ou quantificar séries temporais de novas formas (Zhang et al., 2006; Lai et al., 2010; Verplancke et al., 2010). Um dos avanços mais interessantes surgiu da aplicação da teoria de redes complexas na análise da dinâmica de séries temporais (Campanharo et al., 2011; Zhang & Small, 2006; Lacasa

et al., 2008; Xu. et al., 2008; Gao & Jin, 2009). O método proposto recentemente por Campanharo *et al.* (Campanharo et al., 2011; Campanharo & Ramos, 2016a), capaz de mapear uma série temporal em uma rede complexa, não possui restrições de uso. Ou seja, o mesmo pode ser aplicado em séries temporais estacionárias, não-estacionárias, longas ou mesmo curtas. Além disso, as características de diversas séries experimentais podem ser analisadas com base em um número reduzido de amostras, sob a luz da teoria de redes complexas. É importante ressaltar que o mapeamento proposto por Campanharo *et al.* (Campanharo et al., 2011; Campanharo & Ramos, 2016a) para a análise de séries temporais, e em particular dados de EEG, possui um único parâmetro, a saber, o número de quantis Q . Além disso, análises realizadas em trabalhos anteriores (Campanharo et al., 2011; Campanharo & Ramos, 2016a) mostraram que os resultados produzidos por esse mapeamento são robustos à variação dos valores de Q . Nesse sentido, o objetivo deste trabalho é utilizar esse mapeamento em uma aplicação inédita, mais especificamente, no estudo da dinâmica de séries temporais de EEG de pacientes com a DA, (i) para a distinção entre envelhecimento e a DA e (ii) identificação dos estágios da DA em pacientes doentes.

Neste trabalho todas as simulações foram executadas em um computador Core i3 (4Gb de memória). Convém ressaltar que os *softwares* utilizados para a geração dos resultados são gratuitos. Mais especificamente, os códigos foram escritos nas linguagens de programação *C* e *Python* com os *softwares* *Code::Blocks* (Code::Blocks, 2008) e *Jupyter* (Anaconda, 2011), respectivamente. Os gráficos foram gerados com o *software qtGrace* (Turner, 1995).

1.5 Estrutura do texto

Os demais capítulos do texto estão organizados da seguinte forma:

- Capítulo 2: apresenta um estudo sobre redes complexas, mais especificamente, suas características principais e as medidas utilizadas na caracterização to-

pológica de tais redes.

- Capítulo 3: descreve o mapeamento de uma série temporal em uma rede complexa utilizado neste trabalho, como também a base de dados de pacientes saudáveis e com a DA.
- Capítulo 4: apresenta os resultados obtidos.
- Capítulo 5: apresenta as conclusões deste trabalho e sugestões para trabalhos futuros.

2 REDES COMPLEXAS

Redes complexas são descritas por um conjunto de vértices (nós) e arestas (conexões ou ligações) e algum tipo de interação entre seus vértices (Newman, 2010). Diversos sistemas de importância em neurociência, linguística e social podem ser representados por redes complexas. Exemplos incluem o estudo do funcionamento do cérebro, onde as estruturas cerebrais são os vértices e as arestas representam suas funções; o estudo das relações semânticas entre as palavras, onde as palavras são definidas como os vértices e os seus significados como as arestas; a propagação de notícias falsas (em inglês, fake news) onde as informações são os vértices e as pessoas representam as arestas.

Matematicamente, uma rede complexa $g = (\mathcal{N}, \mathcal{L})$ é constituída de um conjunto de N vértices, $\mathcal{N} = \{n_1, n_2, \dots, n_N\}$ e um conjunto de M arestas, $\mathcal{L} = \{l_1, l_2, \dots, l_M\}$. As arestas podem ser direcionadas - quando o sentido da ligação é considerado, e não-direcionadas - caso contrário. Se as arestas possuem intensidade, a cada aresta é associado um peso, e neste caso, a rede é representada por $g = (\mathcal{N}, \mathcal{L}, \mathcal{W})$, onde $\mathcal{W} = \{w_1, w_2, \dots, w_M\}$ constitui o conjunto de pesos das M arestas. As redes complexas podem ser *estáticas*, quando não há variação do número de vértices, arestas ou na configuração das mesmas; ou *dinâmicas*, quando há variação de pelo menos uma dessas componentes. O tipo mais comum de rede complexa é aquele cujas arestas são direcionadas e com peso (Costa et al., 2007) (Fig. 2).

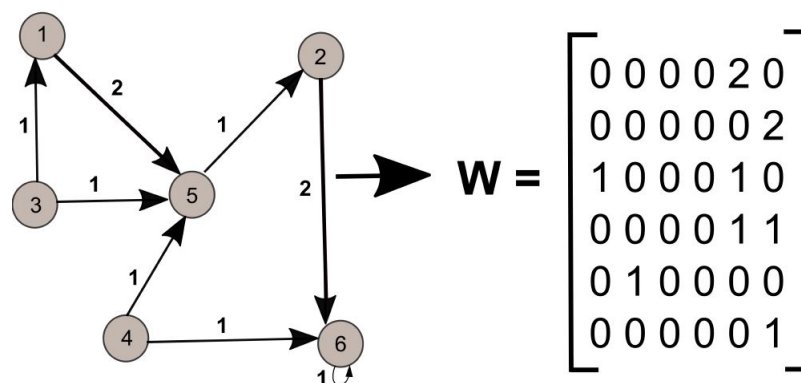


Figura 2: Exemplo de rede direcionada e com peso, e sua matriz de adjacência com peso W .

A partir de uma rede direcionada e com peso, é possível obter suas versões sem peso e sem direção com base nas operações de limiarização e simetrização, respectivamente. A Figura 3 exemplifica o uso de tais operações, onde a limiarização é realizada atribuindo-se peso unitário a todas as arestas da rede e a simetrização é realizada desconsiderando-se a direção entre pares de vértices da rede associada.

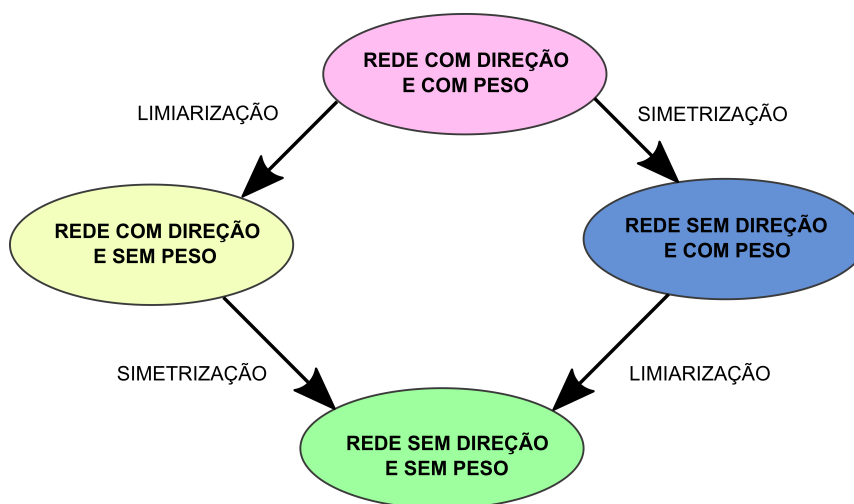


Figura 3: Representação esquemática das relações entre os tipos de redes. Fonte: Adaptada de (Rodrigues, 2007).

Computacionalmente, uma rede complexa pode ser armazenada em dois tipos diferentes de estrutura de dados, a saber, *matriz de adjacência* ou *lista*. Na

representação por *matriz de adjacência*, se dois vértices n_i e n_j estiverem conectados, a posição a_{ij} da matriz de adjacência A será igual a um; caso contrário, será igual a zero. Em uma rede não-direcionada, tem-se que $a_{ij} = a_{ji}$ para $i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$, e nesse caso, A é simétrica. Em uma rede direcionada, as linhas da matriz A representam as ligações que partem do vértice n_i para n_j , e as colunas de A representam as ligações que chegam de n_j por n_i . Em uma rede com peso, a posição w_{ij} da matriz de adjacência com pesos W será igual ao peso da aresta que conecta n_i e n_j . Se a rede não possuir autoconexões, então $a_{ii} = 0$ para $i = 1, 2, \dots, N$. Na representação de uma rede sem peso e sem direção em lista, apenas os pares de vértices (n_i, n_j) conectados por uma aresta são armazenados. Para redes com peso e sem direção, a lista possui um terceiro elemento relacionado ao peso da aresta, ou seja, (n_i, n_j, w_{ij}) .

Cada estrutura de armazenamento tem suas vantagens e desvantagens. O uso de listas permite maior economia de memória para redes esparsas, ou seja, redes que possuem poucos vértices conectados entre si, embora o acesso às ligações não seja um procedimento trivial. O acesso as ligações em matrizes de adjacência é mais simples, porém, este tipo de estrutura de dado é somente bem aproveitado para redes densas, ou seja, redes que possuam a maior parte de seus vértices conectados entre si (Rodrigues, 2007).

2.1 Aspectos históricos

O estudo de redes complexas tem sido de domínio de um ramo da Matemática Discreta, conhecida como *Teoria de Grafos*. A teoria de grafos teve o seu nascimento em 1736, quando o matemático Leonardo Euler resolveu o problema das “sete pontes de Königsberg” (território da Rússia até 1945, atual Kaliningrado), cidade cortada pelo rio Pregel, o qual a dividia em duas ilhas A e D e que, juntas, formavam um complexo que continha sete pontes (Fig.4). Discutia-se a possibilidade de realizar um passeio por tais ilhas, passando uma única vez em todas as sete pontes e voltar ao ponto de origem. Euler transformou os caminhos em arestas e as regiões

de terra em vértices, criando possivelmente o primeiro grafo da história (Newman, 2010). Euler concluiu que tal passeio seria possível caso todos os vértices tivessem um número par de arestas ou que dois vértices tivessem um número ímpar de arestas.

A teoria de redes complexas, além de fazer uso de medidas desenvolvidas pela teoria dos grafos, utiliza conceitos provenientes da mecânica estatística, da física não-linear e de sistemas complexos. Apesar de existirem semelhanças entre a teoria de redes complexas e a teoria dos grafos, a primeira está relacionada com modelagem de processos dinâmicos reais, onde as redes em estudo evoluem com o tempo, modificando suas estruturas.

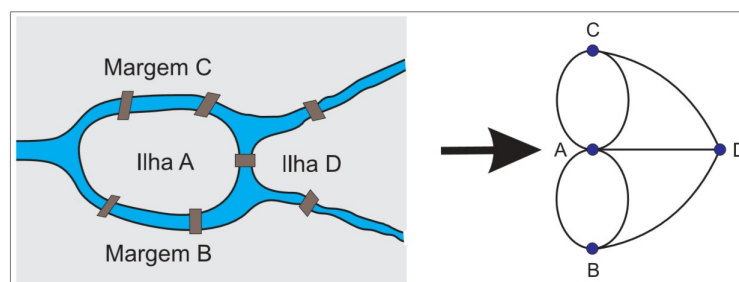


Figura 4: O problema das sete pontes de Königsberg. Representação de Euler por meio de um grafo para o problema. Fonte: (Rodrigues, 2007).

2.2 Propriedades topológicas

A caracterização, análise e discriminação das redes complexas dependem do uso de medidas capazes de expressar suas características topológicas mais relevantes. Algumas destas medidas são: o grau de conectividade, o coeficiente de agrupamento, a modularidade, o comprimento do salto médio e o grau de intermediação.

2.2.1 Grau de conectividade

O grau de conectividade k_i de um vértice n_i é dado pelo número de arestas incidentes desse vértice a outros vértices. Para uma rede não-direcionada,

tem-se

$$k_i = \sum_{j=1}^N a_{ij} = \sum_{j=1}^N a_{ji}, \quad (1)$$

onde a_{ij} ($i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$) são elementos da matriz de adjacência A associada. Para uma rede direcionada, existem dois tipos de graus de conectividade: o grau de conectividade de saída k_i^{out} , igual ao número de arestas que partem do vértice n_i para o vértice n_j , e o grau de conectividade de entrada k_i^{in} , igual ao número de arestas que chegam ao vértice n_i pelo vértice n_j , ou seja,

$$k_i^{out} = \sum_{j=1}^N a_{ij}, \quad (2)$$

$$k_i^{in} = \sum_{j=1}^N a_{ji}, \quad (3)$$

$$k_i = k_i^{out} + k_i^{in}, \quad (4)$$

com $i \neq j$. O grau de conectividade médio $\langle k \rangle$ de uma rede é definido como a média aritmética dos graus de conectividade sobre todos os vértices, ou seja,

$$\langle k \rangle = \frac{1}{N} \sum_{i=1}^N k_i. \quad (5)$$

2.2.2 Força de conectividade

Assim como o grau de conectividade, a força de conectividade s_i de um vértice n_i , em uma rede com peso, é definida como a soma dos pesos das arestas em n_i incidentes a outros vértices n_j . Para redes com peso e sem direção, tem-se que

$$s_i = \sum_{j=1}^N w_{ij} = \sum_{j=1}^N w_{ji}, \quad (6)$$

onde w_{ij} ($i = 1, 2, \dots, N$ e $j = 1, 2, \dots, N$) são elementos da matriz peso W associada. Por outro lado, para redes com peso e com direção, existem dois tipos de força de conectividade: a força de conectividade de saída s_i^{out} , igual ao peso da aresta que parte do vértice n_i para o vértice n_j , e a força de conectividade entrada s_i^{in} , igual

ao peso da aresta que chega ao vértice n_i pelo vértice n_j , ou seja,

$$s_i^{out} = \sum_{j=1}^N w_{ij}, \quad (7)$$

$$s_i^{in} = \sum_{j=1}^N w_{ji}, \quad (8)$$

$$s_i = s_i^{out} + s_i^{in}. \quad (9)$$

A força de conectividade média da rede, é definida como a média aritmética das forças de conectividade sobre todos os vértices, ou seja,

$$\langle s \rangle = \frac{1}{N} \sum_{i=1}^N s_i. \quad (10)$$

2.2.3 Coeficiente de agrupamento médio (C)

O coeficiente de agrupamento é uma medida do grau em que os vértices em uma rede tendem a se agrupar, ou seja, é a tendência de uma rede formar vizinhanças fortemente conectadas. Para uma rede com peso e com direção e um vértice n_i , C_i é definido como a razão entre todos os triângulos com peso e direção que são formados pelo vértice $n_i(t_i)$ e o número total de triângulos possíveis que n_i poderia formar (T_i) (Fagiolo, 2007). Assim,

$$C_i = \frac{t_i}{T_i} = \frac{[\hat{W}_k^{[\frac{1}{3}]} + (\hat{W}_k^T)^{[\frac{1}{3}]}]_{ii}^3}{2[k_i^{tot}(k_i^{tot} - 1) - 2k_i^{\leftrightarrow}]} \quad (11)$$

onde \hat{W}_k é a matriz de pesos normalizada, $\hat{W}_k^{\frac{1}{3}} = (\hat{w}_k)_{ij}^{\frac{1}{3}}$ é a matriz obtida a partir de \hat{W}_k tomando a raiz cúbica de cada elemento, k_i^{tot} é o grau total do vértice n_i , e k_i^{\leftrightarrow} é o número de arestas bilaterais entre o vértice n_i e seus vizinhos. Portanto, o coeficiente de agrupamento médio C da rede é definido como a média aritmética dos C_i para todos os vértices n_i ($i = 1, 2, \dots, N$), ou seja,

$$C = \frac{1}{N} \sum_{i=1}^N C_i. \quad (12)$$

A Figura 5 apresenta uma rede com peso e com direção constituída por 4 vértices. Os valores de t_i , k_i^{tot} , k_i^{\leftrightarrow} e C_i são apresentados na Tabela 1. Desta forma, os valores

dos coeficientes de agrupamento associados aos vértices 1, 2, 3 e 4 são dados por 0,25, 0,50, 0,50 e 0,00, respectivamente. O coeficiente de agrupamento médio dessa rede é dado por $C \approx 0,31$.

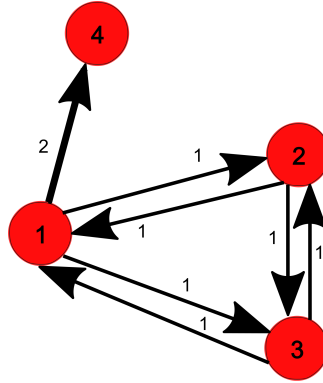


Figura 5: Exemplo de rede com peso e com direção, constituída de 4 vértices. Com base nas Equações 11 e 12, o coeficiente de agrupamento médio dessa rede é dado por $C \approx 0,31$.

Tabela 1: Valores de t_i , k_i^{tot} , $k_i^{↔}$ e C_i para cada um dos vértices da rede da Figura 5.

| Vértices | t_i | k_i^{tot} | $k_i^{↔}$ | C_i |
|----------|-------|-------------|-----------|-------|
| 1 | 8,00 | 5,00 | 2,00 | 0,25 |
| 2 | 8,00 | 4,00 | 2,00 | 0,50 |
| 3 | 8,00 | 4,00 | 2,00 | 0,50 |
| 4 | 0,00 | 1,00 | 0,00 | 0,00 |

2.2.4 Modularidade (M)

Diversas redes reais possuem estruturas modulares, ou seja, um grande número de arestas distribuídas entre vértices de um mesmo módulo, porém um número reduzido de arestas distribuídas entre vértices de módulos distintos. Tais módulos, também conhecidos como comunidades, são formados por vértices que possuem alguma relação de semelhança entre si. Redes de contatos sexuais, redes de citações e redes de amigos são exemplos de redes modulares (Bagnato, 2018).

A modularidade é uma medida que quantifica a possibilidade de se dividir uma rede em módulos (Guimerá et al., 2004). Para uma partição p , a modularidade \mathcal{M}_p pode ser definida por:

$$\mathcal{M}_p = \sum_{s=1}^r \left[\frac{M_s}{M} - \left(\frac{d_s}{2M} \right)^2 \right], \quad (13)$$

onde r é o número de módulos da rede, M é o total de arestas na rede, M_s é o número de arestas entre os vértices pertencentes ao s -ésimo módulo e d_s é a soma dos graus de conectividade dos vértices pertencentes ao s -ésimo módulo. A modularidade \mathcal{M} de uma rede é definida como o máximo valor de \mathcal{M}_p dentre todas as possíveis partições. Quando a rede é composta por um único módulo, então $\mathcal{M}_p = 0$. Como toda rede pode ser particionada dessa forma, conclui-se que $\mathcal{M} \geq 0$. A Figura 6 apresenta uma rede sem peso e sem direção, constituída de 6 vértices e particionada em dois módulos. Nela, os vértices 1, 2 e 3 pertencem ao módulo 1 e os vértices 4, 5 e 6 pertencem ao módulo 2. Utilizando-se a Equação 13, com $r = 2$, $M = 7$, $M_1 = 3$, $d_1 = 7$, $M_2 = 3$, $d_2 = 7$, obtém-se a modularidade de tal rede, dada por $\mathcal{M}_p = 0,35$.

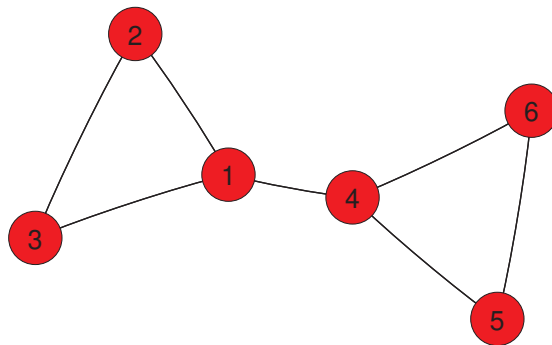


Figura 6: Exemplo de rede modular sem peso e sem direção, constituída de 6 vértices, 7 arestas e 2 módulos. Com base na Equação 13, a modularidade dessa rede é dada por $\mathcal{M}_p = 0,35$.

2.2.5 Comprimento do salto médio (Δ)

A partir da matriz de adjacência A , é possível obter a matriz de transição de Markov P , de tal forma que cada elemento $p_{ij} = a_{ij} / \sum_{j=1}^N a_{ij}$ represente a probabilidade de transição do vértice n_i para o vértice n_j . Assim, é possível realizar um caminho aleatório sobre a rede R e calcular o comprimento do salto médio Δ (Campanharo & Ramos, 2016a), definido como:

$$\Delta = \frac{1}{S} \sum_{s=1}^S \delta_s(i, j), \quad (14)$$

onde $s = S$ são os saltos de comprimentos $\delta_s(i, j) = |i - j|$, onde $i, j = 1, \dots, N$ são os índices dos vértices definidos por P . O cálculo do comprimento do salto médio, pode ser calculado com base na expressão :

$$\Delta = \frac{1}{N} \text{tr}(GP^T), \quad (15)$$

onde P^T é a transposta de P , G é a matriz $N \times N$ com os elementos $g_{i,j} = |i - j|$ e tr é o traço da matriz resultante. O valor de Δ para a rede da Figura 7 pode ser obtido com base nas matrizes:

$$W = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 2 & 0 & 0 \end{bmatrix}, \quad G = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{bmatrix} \quad \text{e} \quad P^T = \begin{bmatrix} 0,50 & 0,00 & 1,00 \\ 0,50 & 0,66 & 0,00 \\ 0,00 & 0,33 & 0,00 \end{bmatrix}$$

e, a partir da Equação 15 o comprimento do salto médio da rede é dado por $\Delta \approx 0,94$.

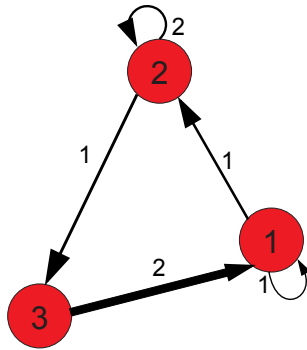


Figura 7: Exemplo de rede com peso e com direção, constituída de 3 vértices. Com base na Equação 15, o comprimento do salto médio dessa rede é dado por $\Delta \approx 0,94$.

2.2.6 Grau de Intermediação (B)

O grau de intermediação (em inglês, *betweenness centrality*) de um dado vértice v mede a importância desse vértice na rede. O mesmo é definido como o somatório da razão de caminhos mais curtos de todos os pares de vértices existentes na rede que passam por um vértice v pelos caminhos mais curtos do vértice s ao vértice t , ou seja,

$$B(v) = \sum_{s \neq v \neq t} \frac{\phi_{st}(v)}{\phi_{st}}, \quad (16)$$

onde ϕ_{st} é o número total de menores caminhos do vértice s para o vértice t e $\phi_{st}(v)$ é o número de menores caminhos que passam por v . Os valores dos graus de intermediação podem ser normalizados e, nesse caso, a Equação 16 deverá ser dividida por $(n - 1)(n - 2)/2$ para redes sem direção e por $(n - 1)(n - 2)$ para redes direcionadas, onde n é o número de vértices da rede. A Figura 8 apresenta um exemplo de uma rede sem peso e sem direção constituída de 5 vértices. A Tabela 2 exemplifica o cálculo do grau de intermediação para o vértice 4 ($B(4)$) para essa rede, dado por $B(4) \approx 0,58$. O cálculo do grau de intermediação para os demais vértices é feito de maneira análoga à apresentada para o vértice 4. Tais valores são dados por, $B(1) \approx 0,08$, $B(2) \approx 0,16$, $B(3) \approx 0,16$ e $B(5) = 0,00$.

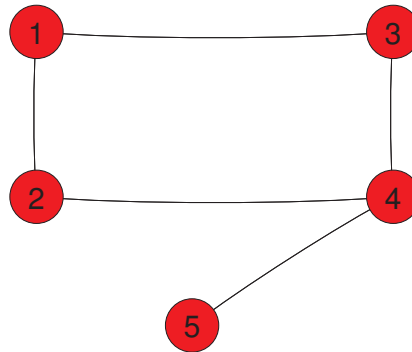


Figura 8: Exemplo de rede sem peso e sem direção, constituída de 5 vértices e 5 arestas. O grau de intermediação para os vértices 1, 2, 3, 4 e 5 são dados por, 0,08, 0,16, 0,16, 0,58 e 0,00, respectivamente. Tal rede possui grau de intermediação médio dado por 0,58.

Tabela 2: Cálculo do grau de intermediação para o vértice 4.

| Vértices | $\phi_{st}(4)$ | ϕ_{st} | $\frac{\phi_{st}(4)}{\phi_{st}}$ |
|----------|----------------|-------------|----------------------------------|
| {1,5} | 2,00 | 2,00 | 1,00 |
| {1,3} | 0,00 | 1,00 | 0,00 |
| {1,2} | 0,00 | 1,00 | 0,00 |
| {2,5} | 1,00 | 1,00 | 1,00 |
| {2,3} | 1,00 | 2,00 | 0,50 |
| {3,5} | 1,00 | 1,00 | 1,00 |

As complexidades computacionais dos algoritmos utilizados para o cálculo da modularidade, do coeficiente de agrupamento médio, do comprimento do salto médio e do grau de intermediação médio são da ordem de n^3 , n^3 , n e $nm + n^2 \log n$, respectivamente (Newman & Girvan, 2004; Fagiolo, 2007; Golub & Van Loan, 1989; Brandes, 2001), onde n é o número de vértices e m é o número de arestas da rede.

3 METODOLOGIA

3.1 Mapeamento de séries temporais em Redes Complexas

Campanharo *et al.* (Campanharo et al., 2011; Campanharo & Ramos, 2016a) propuseram um mapeamento de uma série temporal em uma rede complexa que permite analisar a dinâmica de uma série temporal por meio de um conjunto extenso de propriedades topológicas da rede complexa associada. Dada uma série temporal, primeiramente seus Q quantis são identificados, e então, cada quantil q_i é associado a um vértice $n_i \in \mathcal{N}$ na rede correspondente. Dois vértices n_i e n_j estarão conectados com a aresta $(n_i, n_j, w_{ij}^k) \in \mathcal{L}$, onde o peso w_{ij}^k de cada aresta é dado pelo número de vezes que um dado ponto x_t no quantil q_i é seguido por um ponto x_{t+k} no quantil q_j , para $t = 1, 2, \dots, T - k$ e $k = 1, \dots, k_{max} < T$. A rede complexa resultante possui peso, direção, $Q \approx 2T^{1/3}$ vértices e é conexa, ou seja, não possui vértices isolados (Campanharo et al., 2018).

A Figura 9 ilustra a aplicação desse mapeamento para uma série com $T = 15$ pontos, $Q = 5$ quantis e $k = 1, 2$ e 3 , mapeada em três redes com cinco vértices cada. Os intervalos dos quantis são dados por $[x(0), x(3)[$, $[x(3), x(6)[$, $[x(6), x(9)[$, $[x(9), x(12)[$ e $[x(12), x(15)]$ para os dados ordenados, isto é, $[-8, 740, -4, 912[$, $[-4, 912, 0, 128[$, $[0, 128, 2, 743[$, $[2, 743, 5, 000[$ e $[5, 000, 6, 571]$. Cada quantil é associado a um vértice na rede correspondente. O peso de cada aresta w_{ij}^k representa o número de transições entre os quantis q_i e q_j para um dado valor de k . Os pesos das arestas são dados por $(1, 1, 1)$, $(1, 4, 1)$, $(1, 5, 1)$, $(2, 4, 1)$, $(2, 5, 2)$, $(3, 1, 1)$, $(3, 2, 1)$, $(3, 3, 1)$, $(4, 2, 1)$, $(4, 3, 1)$, $(5, 1, 1)$, $(5, 2, 1)$ e $(5, 4, 1)$ para $k = 1$; $(1, 4, 2)$, $(2, 1, 1)$, $(2, 2, 1)$, $(2, 3, 1)$, $(3, 1, 1)$, $(3, 5, 2)$, $(4, 2, 1)$, $(4, 4, 1)$, $(5, 1, 1)$,

(5, 2, 1) e (5, 5, 1) para $k = 2$ e (1, 2, 1), (2, 1, 1), (2, 2, 1), (2, 5, 1), (3, 2, 1), (3, 4, 1), (3, 5, 1), (4, 3, 1), (4, 5, 1), (5, 1, 1) e (5, 4, 2) para $k = 3$. Transições repetidas entre os quantis resultam em arestas na rede com pesos maiores (representadas na rede correspondente por linhas mais grossas).

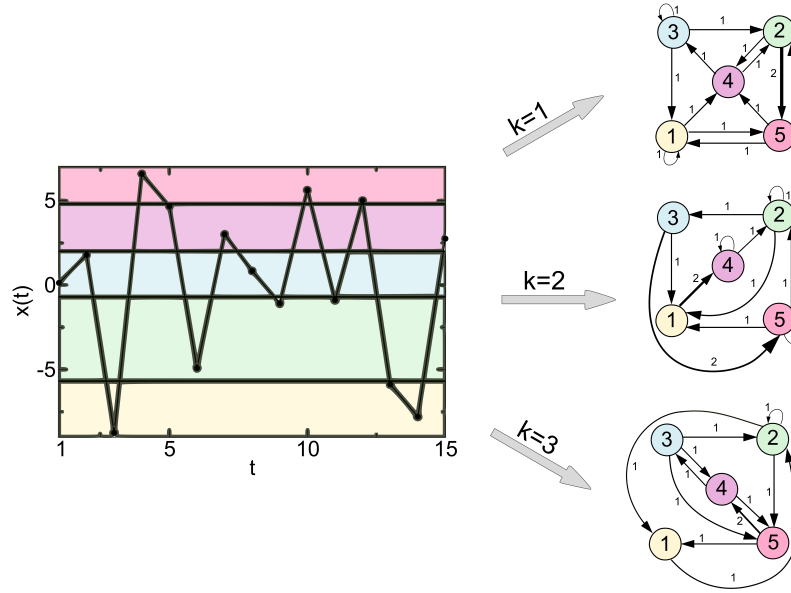


Figura 9: Ilustração do mapeamento proposto por Campanharo *et al.* (Campanharo *et al.*, 2011; Campanharo & Ramos, 2016a) para uma série temporal com $T = 15$ pontos, $Q = 5$ quantis e $k = 1, 2$ e 3 , mapeada em três redes com cinco vértices cada. Observe que, independente do valor de k escolhido, tal mapeamento produz redes com peso, direção e conexas. Os intervalos dos quantis são dados por $[x(0), x(3)[$, $[x(3), x(6)[$, $[x(6), x(9)[$, $[x(9), x(12)[$ e $[x(12), x(15)]$ para os dados ordenados, isto é, $[-8, 740, -4, 912[$, $[-4, 912, 0, 128[$, $[0, 128, 2, 743[$, $[2, 743, 5, 000[$ e $[5, 000, 6, 571]$. Fonte: (Pineda *et al.*, 2019).

O mapeamento em questão foi aplicado em um conjunto de séries temporais sintéticas com características periódicas, pseudo-periódicas, aleatórias e fractais (Campanharo *et al.*, 2011; Campanharo & Ramos, 2016a,b,a). Observou-se que séries temporais com dinâmicas distintas foram mapeadas em redes complexas com topologias também distintas, o que atesta a eficiência do mapeamento proposto. Além disso, tal mapeamento foi utilizado na classificação de dados de EEG de indivíduos saudáveis e com epilepsia (Campanharo *et al.*, 2017, 2018). Mais especi-

ficamente, de dados de EEG medidos na superfície do couro cabeludo de indivíduos (A) sadios com olhos fechados e (B) sadios com olhos abertos e EEG's intracranianos de pacientes com epilepsia (C) dentro e (D) fora da zona epileptogênica, bem como, (E) de pacientes com epilepsia durante o período ictal. Análises feitas mostraram que o mapeamento em estudo foi capaz não só de diferenciar pacientes sadios de pacientes doentes, mas também de distinguir os diferentes padrões (estágios) de uma crise epiléptica, tais como o inter-ictal (alterações nos dados de EEG antes ou depois da crise) e ictal (alterações nos dados de EEG durante a crise). Tais resultados mostraram a eficácia do mapeamento utilizado na distinção de estruturas em dados fisiológicos.

É conhecido na literatura que as séries de EEG de pacientes com a DA apresentam uma diminuição das componentes de alta frequência e um aumento das componentes de baixa frequência, quando comparadas às séries de EEG de pacientes sadios (Ahmadlou et al., 2010). Com base nessa observação, espera-se que o mapeamento em questão seja capaz de diferenciar pacientes sadios de pacientes com a DA. É também conhecido na literatura que os estágios iniciais da DA são caracterizados pelo aumento da atividade das ondas teta (frequências de 4 à 8 Hz) e pela diminuição da atividade das ondas beta (frequências de 13 à 30 Hz). E também que, à medida que a doença progride, tais mudanças são acompanhadas pela diminuição das ondas alfa (frequências de 8 à 13 Hz). Um aumento das ondas delta (frequências de 0,5 à 4 Hz) ocorre no estágio mais avançado da DA (Adeli et al., 2005). Por fim, é conhecido na literatura que a primeira região do cérebro afetada pela DA é o hipocampo e em seguida a doença difunde para as demais regiões cerebrais. Com base em tais informações, esperamos que o mapeamento em questão seja capaz de identificar os estágios da DA de cada um dos pacientes doentes em estudo.

3.2 Análise de Alzheimer via dados de EEG

O estudo e o diagnóstico da DA foram feitos com base em dados de EEG fornecidos pelo Departamento de Física da Universidade do Estado da Flórida

(EUA). Tal base de dados é constituída de 19 canais de EEG de 24 idosos saudáveis (grupo de controle), sem históricos de distúrbios neurológicos ou psiquiátricos, com idade média de 71 anos (variação entre 61 e 83 anos). Essa base de dados compreende, ainda, 160 pacientes diagnosticados com a DA pelo Instituto Nacional de Distúrbios Neurológicos, Desordens Comunicativas e Acidente Vascular Encefálico (NINCDS-ADRDA), com idade média de 74 anos (variação entre 53 e 85 anos). Os dados foram coletados em pacientes com olhos abertos e olhos fechados. A duração dos trechos analisados foi de oito segundos, a uma taxa de amostragem de 128 Hz, totalizando séries com 1.024 pontos para cada um dos indivíduos e para cada um dos 19 canais disponíveis. Neste trabalho, dividimos esse banco de dados em quatro conjuntos (denotados A, B, C e D) com 12 pacientes cada. O conjunto A compreende pacientes saudáveis com os olhos abertos, o conjunto B pacientes saudáveis com os olhos fechados, o conjunto C pacientes doentes com olhos abertos e, por fim, o conjunto D pacientes doentes com os olhos fechados. Convém ressaltar que tal base de dados não disponibiliza o estágio da DA que cada paciente se encontra.

4 RESULTADOS

4.1 Distinção entre envelhecimento e a DA

O mapeamento em estudo foi aplicado ao problema de diferenciação entre pacientes saudáveis (olhos abertos (A) e olhos fechados (B)) e pacientes doentes (olhos abertos (C) e olhos fechados (D)). Para isso, foram utilizadas séries de EEGs relativas aos canais F_7 , F_8 e P_z . A escolha de tais canais se justifica pelo fato dos canais F_7 e F_8 serem canais mais próximos do hipocampo - que corresponde à primeira região do cérebro a ser afetada pela DA e P_z ser o canal próximo ao lobo parietal - que constitui um canal selecionado aleatoriamente para as análises subsequentes.

Como cada série temporal possui $T = 1.024$ pontos, foram utilizados $Q = 2(1.024)^{\frac{1}{3}} \approx 20$ quantis (ou vértices) e $k = 1, 2, \dots, 25$ em todas as simulações. Assim, 48 sinais de EEG foram mapeados em 1.200 grafos de quantis (ou 1.200 matrizes de adjacência A_k), e portanto, foram obtidas 1.200 matrizes de transição de Markov W_k com $Q^2 = 400$ elementos cada. As Figuras 10, 13 e 16 apresentam exemplos de sinais de EEG de indivíduos de cada um dos conjuntos A, B, C e D e para os três canais em estudo. Note que, em todos os casos, os EEGs de pacientes saudáveis diferem dos EEGs de pacientes doentes.

Para os conjuntos A, B, C e D e para os canais F_7 , F_8 e P_z , foram calculados $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k usando as Equações 11, 13, 15 e 16, respectivamente (Figs. 11, 14 e 17). Observe que, em todos os casos, as curvas associadas à pacientes sadios (A e B) e com à DA (C e D) formam dois clusters distintos com máxima separação para um dado valor de k . A Tabela 3 apresenta os

valores de k , correspondentes à maior distinção entre pacientes saudáveis e com a DA com base nas curvas de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , independentemente da condição “olhos abertos” e “olhos fechados”, relativos aos três canais em estudo.

Tabela 3: Valores de k escolhidos para cada canal e cada classificador de rede em estudo.

| | C | M | Δ | B |
|-------|----------|----------|----------|----------|
| F_7 | 9 | 6 | 10 | 6 |
| F_8 | 7 | 5 | 8 | 5 |
| P_z | 12 | 13 | 11 | 12 |

As Figuras 12, 15 e 18 exibem os boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ para os valores de k apresentados na Tabela 3, para os conjuntos A, B, C e D e canais F_7 , F_8 e P_z , respectivamente. Cada bloxpot indica (de baixo para cima) o menor valor, o primeiro quartil, a mediana, o terceiro quartil e o maior valor de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, respectivamente. Observe que o mapeamento utilizado permite uma discriminação entre os pacientes saudáveis (A e B) e pacientes doentes (C e D), independentemente do caracterizador de rede e canal escolhidos. Além disso, para verificar as diferenças entre as médias dos conjuntos A, B, C e D foi utilizado o teste estatístico ANOVA (Âpendice A). Concluiu-se que há uma diferença significativa entre as médias correspondentes para um intervalo de confiança (IC) de 95% e um valor-p inferior a 0,05. Para as condições “olhos abertos” e “olhos fechados” não houve evidências de diferença, já para as condições “saudáveis” e “doentes” houve evidências de diferença (Tabelas 4, 6 e 8). Em seguida, o cálculo da área sob a curva ROC (A_{ROC}) (Âpendice A) foi utilizado com a finalidade de quantificar com precisão a diferença entre os grupos de pacientes sadios e doentes. As Tabelas 5, 7 e 9 apresentam os valores de A_{ROC} referentes aos caracterizadores C , M , Δ e B para a distinção entre pacientes saudáveis e doentes, que correspondem às combinações entre os conjuntos A e C, B e D, A e D e B e C, respectivamente. Com base nos caracterizadores utilizados na discriminação entre pacientes sadios e doentes, Δ exibe

os melhores resultados com valores para a curva ROC iguais a 1.

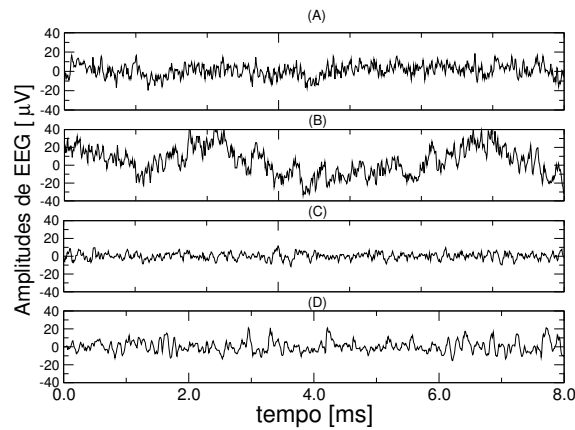


Figura 10: Exemplos de séries de EEG para cada um dos conjuntos em estudo (canal F_7 - lobo frontal). De cima para baixo: A (paciente saudável, olhos abertos), B (paciente saudável, olhos fechados), C (paciente doente, olhos abertos) e D (paciente doente, olhos fechados).

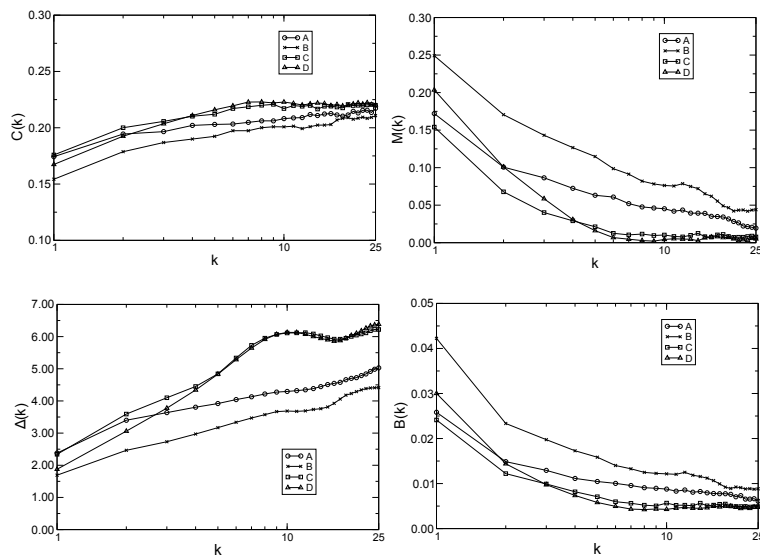


Figura 11: $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , (canal F_7 - lobo frontal) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos), B (pacientes saudáveis, olhos fechados), C (pacientes doentes, olhos abertos) e D (pacientes doentes, olhos fechados).

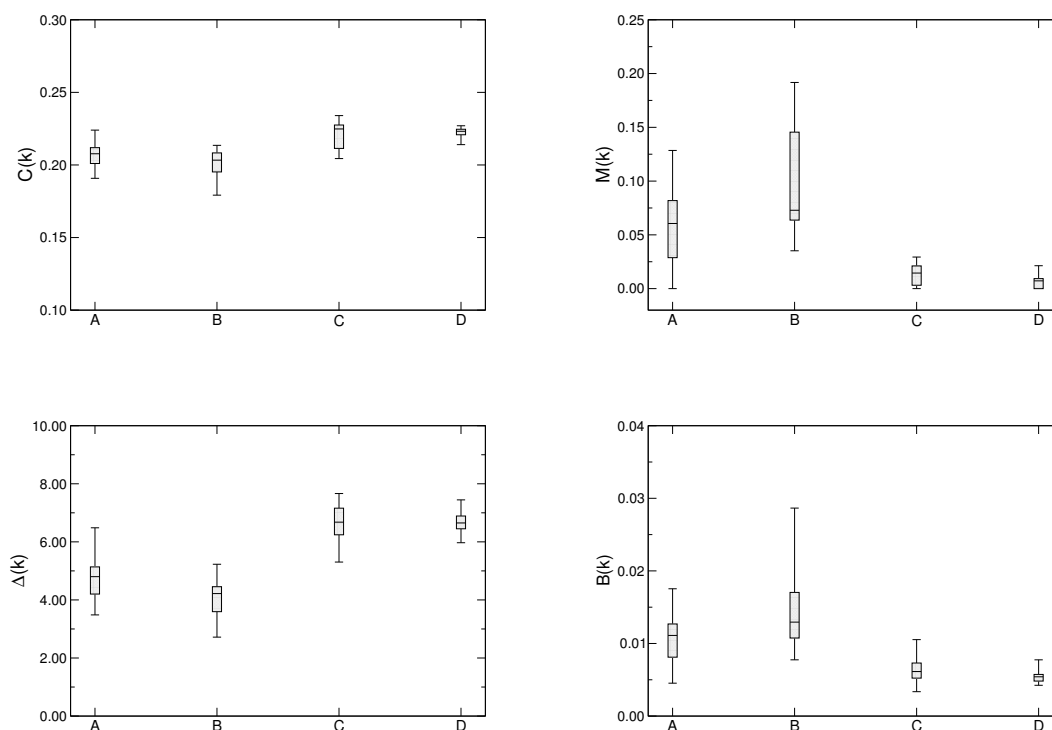


Figura 12: Boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, (canal F_7 - lobo frontal) para os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados) e $k = 9, 6, 10$ e 6 , respectivamente. Pacientes em diferentes condições de saúde possuem medianas distintas (linhas centrais de cada caixa), dadas por 0,2077, 0,2033, 0,2248 e 0,2232 para $C(k)$; 0,0605, 0,0728, 0,0145 e 0,0072 para $M(k)$; 4,8010, 4,2190, 6,6790 e 6,6510 para $\Delta(k)$ e 0,0111, 0,0129, 0,0061 e 0,0054 para $B(k)$, respectivamente.

Tabela 4: Comparação estatística entre as médias de C , M , Δ e $B(k)$ para $k = 9, 6$ e 10 e 6 , (canal F_7 - lobo frontal) respectivamente entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados), com base no teste estatístico ANOVA.

| | C | M | Δ | B |
|------------------|-------------------------|-------------------------|----------------------------|-------------------------|
| IC_{SD}^1 | [-0,0221; -0,0071] | [0,0200; 0,0751] | [-2,5075; -1,3472] | [0,0012; 0,0078] |
| <i>valor - p</i> | $2,3900 \times 10^{-8}$ | $5,5930 \times 10^{-9}$ | $1,9000 \times 10^{-14}$ | $2,3360 \times 10^{-7}$ |

¹S-saudáveis; D-doentes.

Tabela 5: Áreas sob as curvas ROC de C , M , Δ e B para $k = 9, 6, 10$ e 6 (canal F_7 - lobo frontal) entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados).

| | C | M | Δ | B |
|---------------|----------|----------|----------------------------|----------|
| $A_{ROC}(AC)$ | 0,8472 | 0,8785 | 0,9653 | 0,8368 |
| $A_{ROC}(BD)$ | 1,0000 | 1,0000 | 1,0000 | 0,9965 |
| $A_{ROC}(AD)$ | 0,9375 | 0,9236 | 0,9653 | 0,9167 |
| $A_{ROC}(BC)$ | 0,9167 | 1,0000 | 1,0000 | 0,9792 |

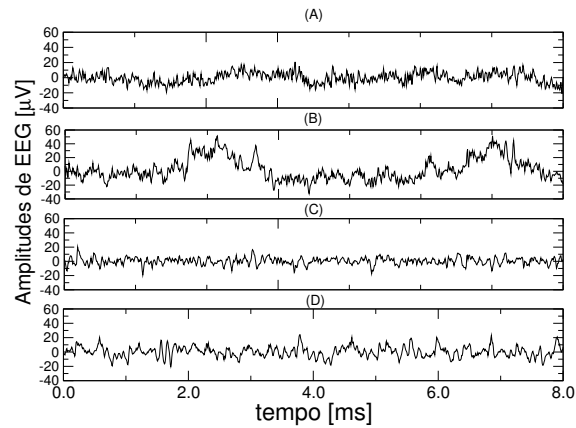


Figura 13: Exemplos de séries de EEG para cada um dos conjuntos em estudo (canal F_8 - lobo frontal). De cima para baixo: A (paciente saudável, olhos abertos), B (paciente saudável, olhos fechados), C (paciente doente, olhos abertos), D (paciente doente, olhos fechados).

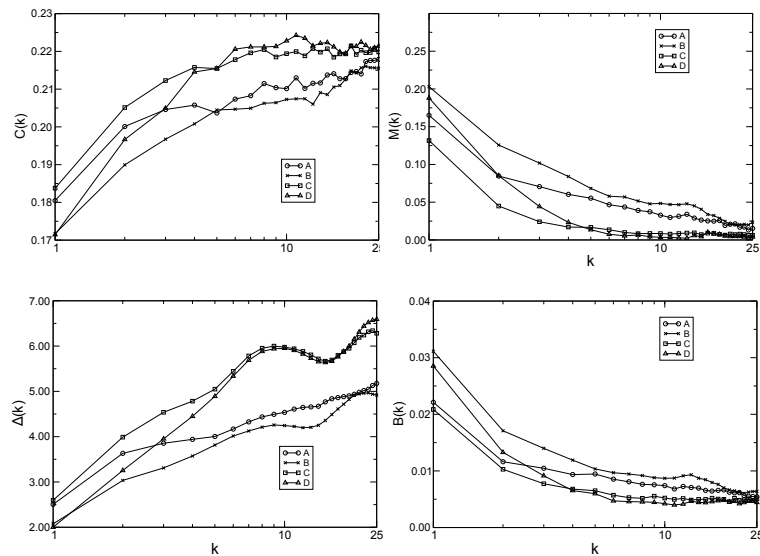


Figura 14: $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , (canal F_8 - lobo frontal) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos), B (pacientes saudáveis, olhos fechados), C (pacientes doentes, olhos abertos) e D (pacientes doentes, olhos fechados).

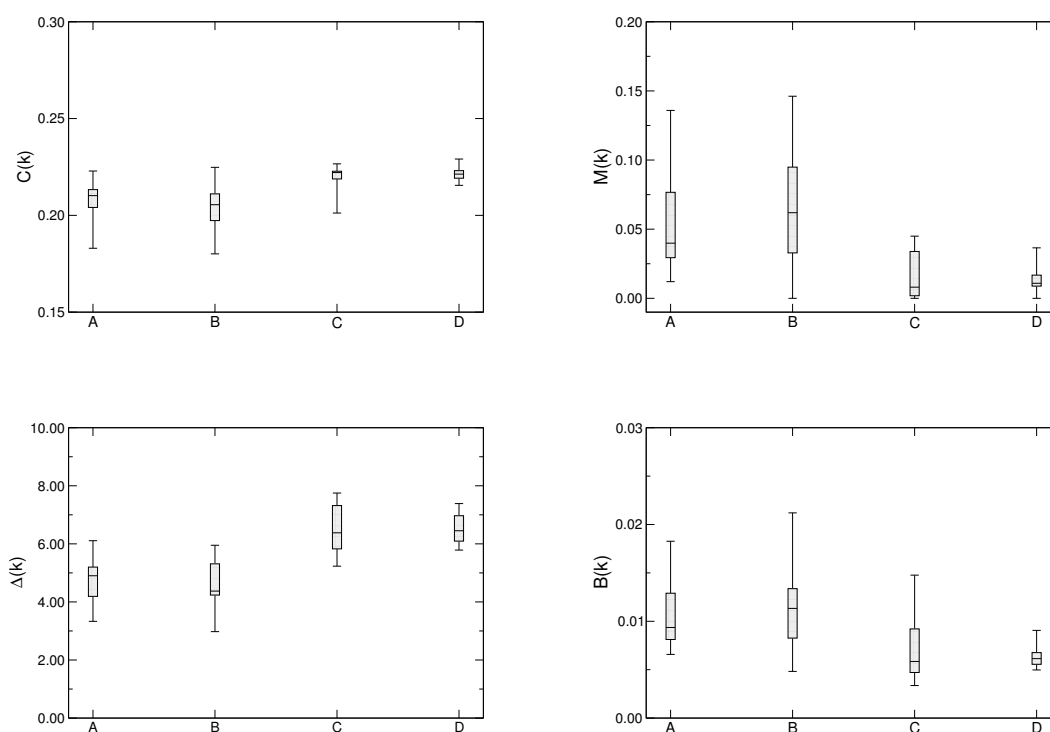


Figura 15: Boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, (canal F_8 - lobo frontal) para os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados) e $k = 7, 5, 8$ e 5 , respectivamente. Pacientes em diferentes condições de saúde possuem medianas distintas (linhas centrais de cada caixa), dadas por 0,2102, 0,2055, 0,2221 e 0,2213 para $C(k)$; 0,0398, 0,0619, 0,0080 e 0,0108 para $M(k)$; 4,9010, 4,3730, 6,3800 e 6,4520 para $\Delta(k)$ e 0,0093, 0,0113, 0,0058 e 0,0061 para $B(k)$, respectivamente.

Tabela 6: Comparação estatística entre as médias de C , M , Δ e B para $k = 7, 5, 8$ e 5 , (canal F_8 - lobo frontal) respectivamente entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados), com base no teste estatístico ANOVA.

| | C | M | Δ | B |
|------------------|-------------------------|-------------------------|----------------------------|------------------|
| IC_{SD}^2 | [-0,0186; -0,0041] | [0,0118; 0,0653] | [-2,3715; -1,0766] | [0,0006; 0,0065] |
| <i>valor - p</i> | $2,3860 \times 10^{-6}$ | $1,0470 \times 10^{-5}$ | $3,3750 \times 10^{-10}$ | 0,000129 |

²S-saudáveis; D-doentes.

Tabela 7: Áreas sob as curvas ROC de C , M , Δ e B para $k = 7, 5, 8$ e 5 (canal F_8 - lobo frontal) entre os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados).

| | C | M | Δ | B |
|---------------|----------|----------|----------------------------|----------|
| $A_{ROC}(AC)$ | 0,8472 | 0,8194 | 0,9306 | 0,7917 |
| $A_{ROC}(BD)$ | 0,9028 | 0,8438 | 0,9861 | 0,8194 |
| $A_{ROC}(AD)$ | 0,9097 | 0,9236 | 0,9792 | 0,9097 |
| $A_{ROC}(BC)$ | 0,8611 | 0,8368 | 0,9375 | 0,7674 |

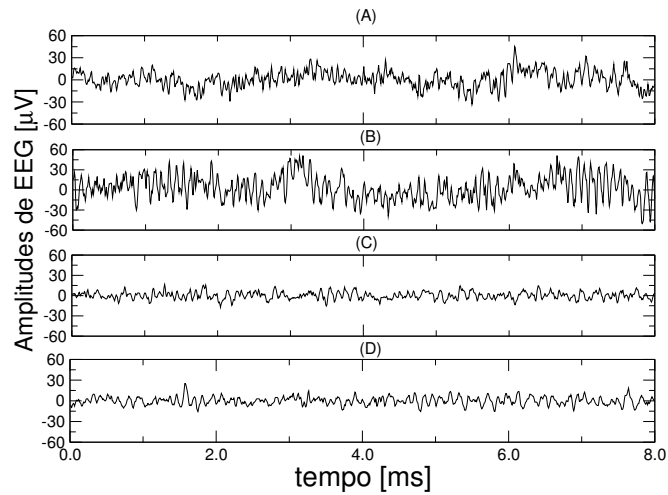


Figura 16: Exemplos de séries de EEG para cada um dos conjuntos em estudo (canal P_z - lobo parietal). De cima para baixo: A (paciente saudável, olhos abertos), B (paciente saudável, olhos fechados), C (paciente doente, olhos abertos), D (paciente doente, olhos fechados).

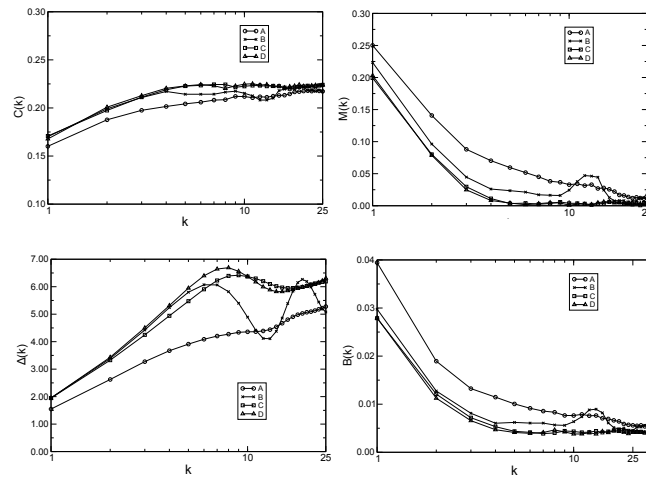


Figura 17: $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k , (canal P_z - lobo parietal) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos), B (pacientes saudáveis, olhos fechados), C (pacientes doentes, olhos abertos) e D (pacientes doentes, olhos fechados).

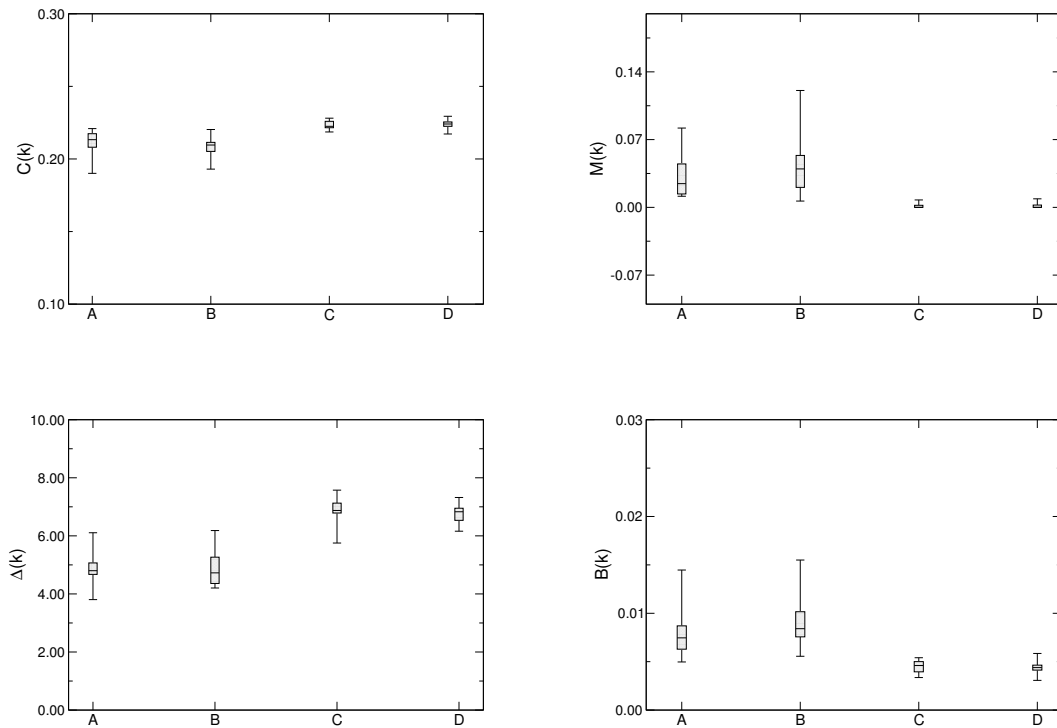


Figura 18: Boxplots de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$, (canal P_z - lobo parietal) para os conjuntos A (saudáveis, olhos abertos), B (saudáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados) e $k = 12, 13, 11$ e 12 , respectivamente. Pacientes em diferentes condições de saúde possuem medianas distintas (linhas centrais de cada caixa), dadas por 0,2133, 0,2096, 0,2225 e 0,2241 para $C(k)$; 0,0245, 0,0397, 0,0000 e 0,0000 para $M(k)$; 4,8000, 4,7260, 6,8750 e 6,8300 para $\Delta(k)$ e 0,0074, 0,0084, 0,0046 e 0,0043 para $B(k)$, respectivamente.

Tabela 8: Comparação estatística entre as médias de C , M , Δ e B para $k = 12, 13, 11$ e 12 , (canal P_z - lobo parietal) respectivamente entre os conjuntos A (saúdáveis, olhos abertos), B (saúdáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados), com base no teste estatístico ANOVA.

| | C | M | Δ | B |
|------------------|--------------------------|-------------------------|----------------------------|-------------------------|
| IC_{SD}^3 | [-0,0165; -0,0067] | [0,0141; 0,0487] | [-2,4256; -1,5763] | [0,0019; 0,0050] |
| <i>valor - p</i> | $4,6230 \times 10^{-10}$ | $1,4820 \times 10^{-7}$ | $2,0000 \times 10^{-16}$ | $2,0250 \times 10^{-9}$ |

³S-saúdáveis; D-doentes.

Tabela 9: Áreas sob as curvas ROC de C , M , Δ e B para $k = 12, 13, 11$ e 12 (canal P_z - lobo parietal) entre os conjuntos A (saúdáveis, olhos abertos), B (saúdáveis, olhos fechados), C (doentes, olhos abertos) e D (doentes, olhos fechados).

| | C | M | Δ | B |
|---------------|----------|----------|----------------------------|----------|
| $A_{ROC}(AC)$ | 0,9722 | 1,0000 | 0,9931 | 0,9757 |
| $A_{ROC}(BD)$ | 0,9792 | 0,9861 | 0,9861 | 0,9931 |
| $A_{ROC}(AD)$ | 0,9653 | 1,0000 | 1,0000 | 0,9826 |
| $A_{ROC}(BC)$ | 0,9931 | 0,9931 | 0,9931 | 1,0000 |

Para fins ilustrativos, a Figura 19 exibe as matrizes de transição de Markov das medianas dos conjuntos A, B, C e D, para $k = 11$ e o canal P_z , sendo esse o canal que melhor diferenciou pacientes saúdáveis de pacientes com a DA. É evidente que séries temporais com dinâmicas distintas correspondem W_{11} com estruturas distintas. Mais especificamente, as matrizes associadas aos pacientes sadios (Fig. 19 (a) e (b)) possuem os pesos mais elevados distribuídos em suas respectivas diagonais principais. À diminuição das atividades das ondas alfa e o aumento das ondas delta fazem com que os pesos nas matrizes correspondentes se distribuam de maneira uniforme nas matrizes de transição de Markov associadas (Fig. 19 (c) e (d)).

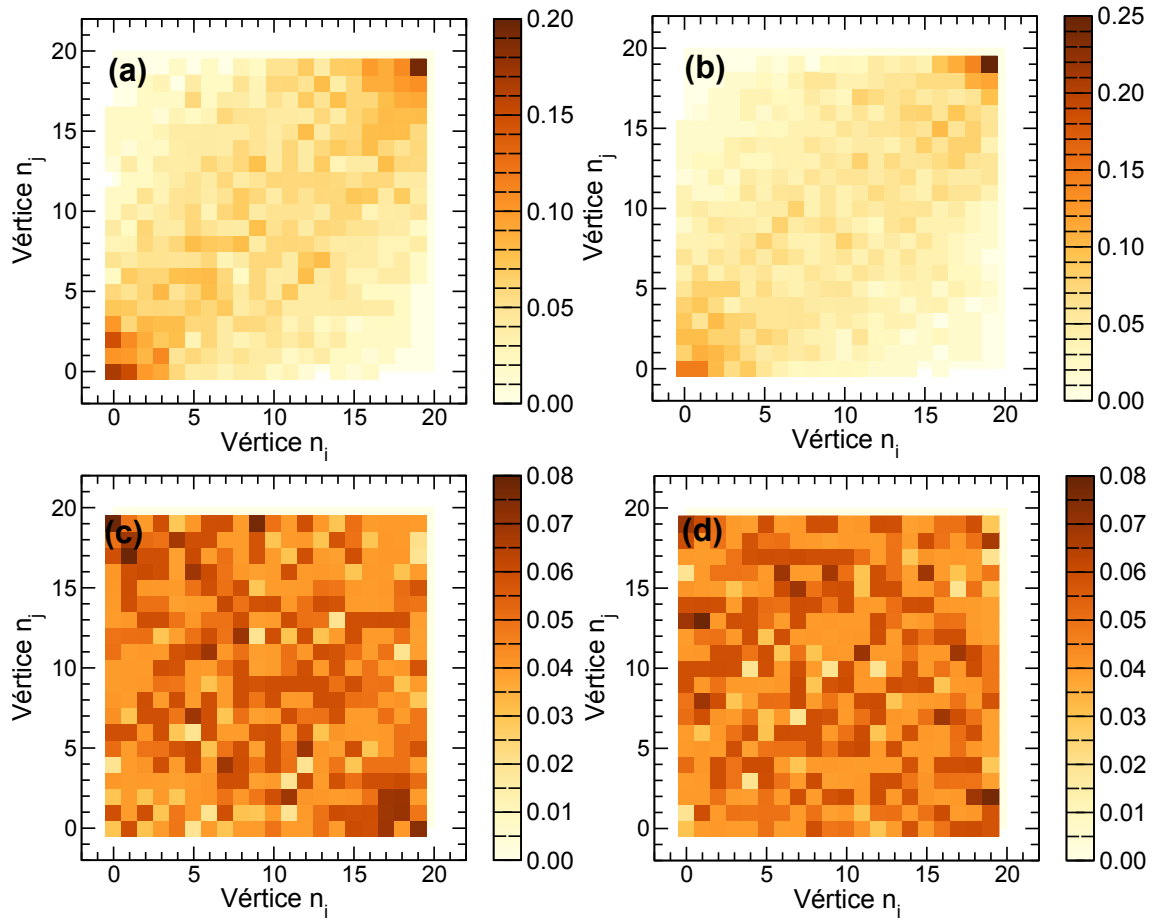


Figura 19: Matrizes de transição de Markov para $k = 11$ (canal P_z - lobo parietal) dos conjuntos (a) (saudáveis, olhos abertos), (b) (saudáveis, olhos fechados), (c) (doentes, olhos abertos) e (d) (doentes, olhos fechados).

4.2 Desempenho dos caracterizadores com base nos canais de EEG

Com base nos 19 canais de EEG, nesta seção foi feito um estudo do desempenho dos caracterizadores de redes complexas já apresentados para a distinção entre pacientes sadios e com a DA (conjuntos A e C, B e D, A e D e B e C), para a distinção entre pacientes sadios e com a DA com olhos abertos (conjuntos A e C) e para a distinção entre pacientes sadios e com a DA com olhos fechados (conjuntos B e D). Nesse sentido, para um dado caracterizador, um dado canal e as combinações dos conjuntos A e C, B e D, A e D e B e C (Fig. 20), A e C (Fig. 21) e B e D (Fig. 22) foi calculado a média dos valores de A_{ROC} , denotado aqui por \hat{A}_{ROC} , os valores de A_{ROC} e os valores de A_{ROC} , respectivamente. Portanto, identificamos o canal de EEG que diferenciou com maior precisão os pacientes saudáveis dos pacientes com a DA. De forma geral, o mapeamento em estudo e, conseqüentemente, os caracterizadores utilizados mostraram-se eficazes na diferenciação de pacientes sadios e com a DA. O comprimento do salto médio (Δ) mostrou-se o caracterizador topológico mais eficaz, possuindo valores da A_{ROC} igual a 1 em todas as análises. Observamos ainda que, independente do caracterizador utilizado o canal P_3 apresentou o melhor desempenho para a diferenciação entre pacientes saudáveis e com a DA. É conhecido que a primeira região do cérebro afetada pela doença é o hipocampo e, em seguida, a doença se difunde para as demais regiões do cérebro. Com base nas análises feitas, concluímos que o canal P_3 obteve o melhor desempenho uma vez que, a base de dados utilizada é composta por pacientes que já não estão mais no estágio inicial da DA.

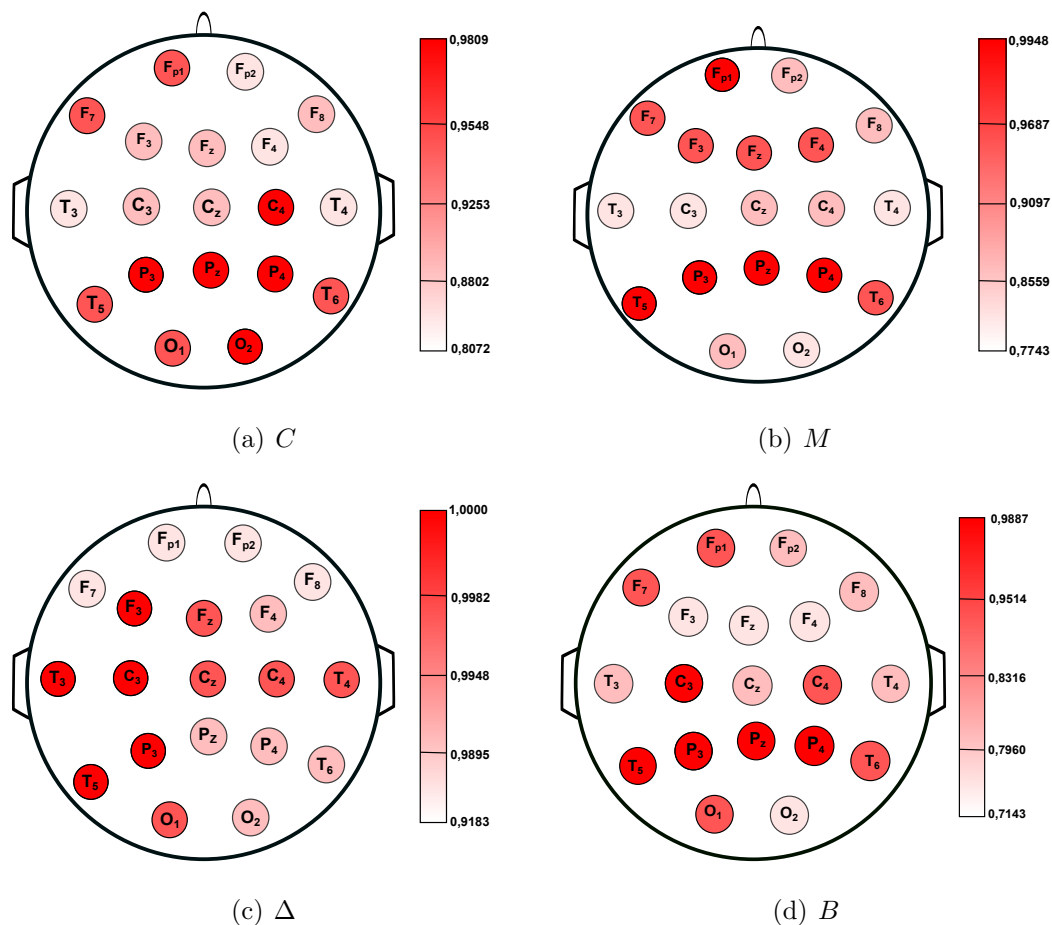


Figura 20: Desempenho dos caracterizadores C , M , Δ e B com base no valores de \hat{A}_{ROC} para os 19 canais de EEG, na distinção entre pacientes sadios e doentes. Com base nas combinações entre os conjuntos A e C, B e D, A e D e B e C, foi calculado a média dos valores de \hat{A}_{ROC} que varia de 0,7143 a 1,0000.

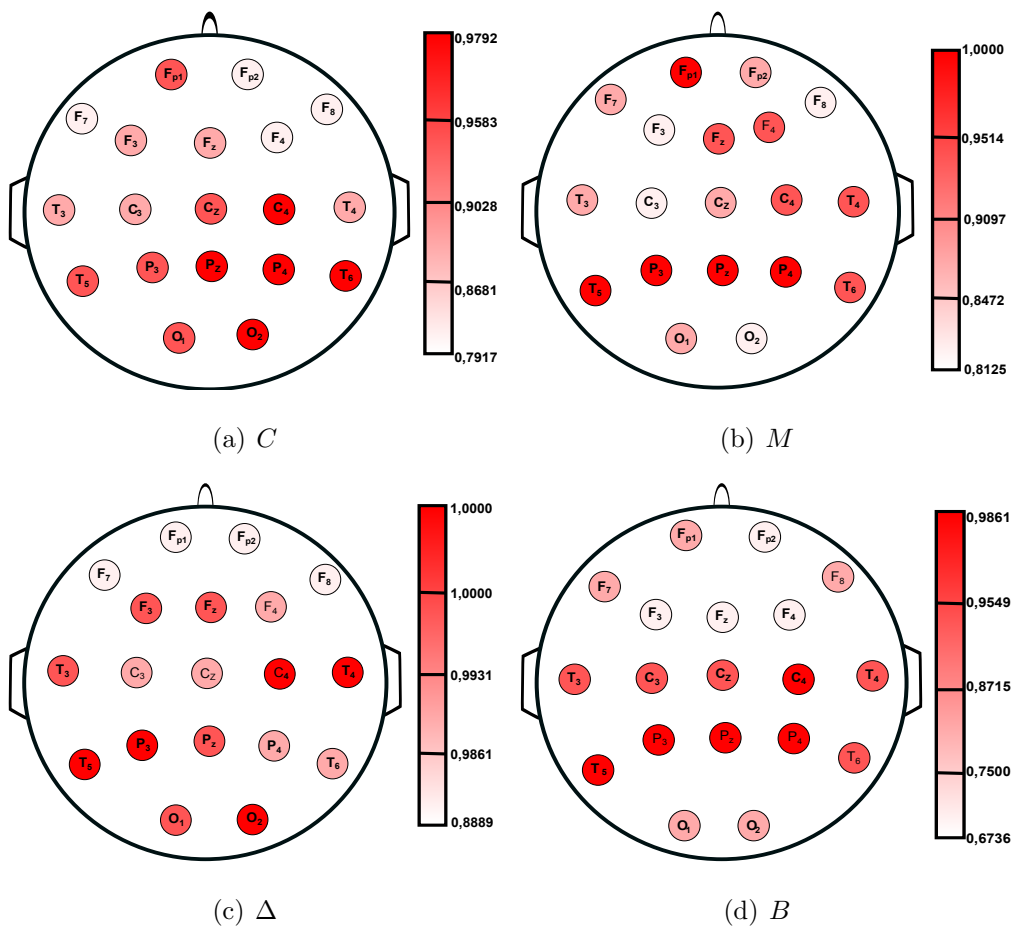


Figura 21: Desempenho dos caracterizadores C , M , Δ e B com base nos valores de A_{ROC} para os 19 canais de EEG, na distinção entre pacientes sadios e doentes com olhos abertos. Com base nas combinações entre os conjuntos A e C, foi calculado os valores de A_{ROC} que varia de 0,6736 a 1,0000.

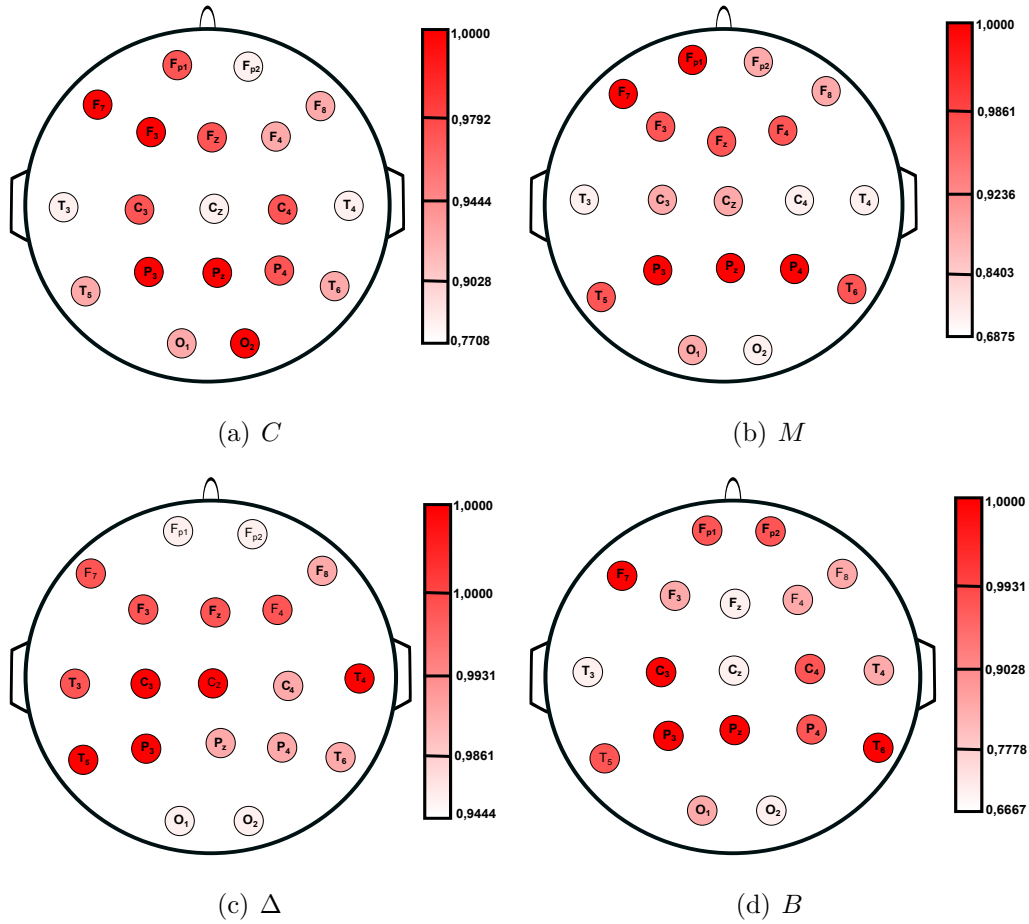


Figura 22: Desempenho dos caracterizadores C , M , Δ e B com base nos valores de A_{ROC} para os 19 canais de EEG, na distinção entre pacientes saudáveis e doentes com olhos fechados. Com base nas combinações entre os conjuntos B e D, foi calculado os valores de A_{ROC} que varia de 0,6667 a 1,0000.

4.2.1 Investigação dos possíveis estágios da DA

Com o objetivo de estudar os estágios da DA dos pacientes doentes com olhos abertos (conjunto C) foi feita uma análise da decomposição dos sinais de EEG em bandas de frequência para o canal P_3 , sendo esse o canal que apresentou o melhor desempenho para a distinção entre pacientes saudáveis e pacientes doentes. Inicialmente, decompomos os sinais de EEG em suas principais ondas cerebrais, a saber, teta, beta, alfa e delta. Tal decomposição foi realizada a partir da função

transformada de *Wavelet* (Daubechies, 1990; Higor B. Santos & Stedile, 2014) e para os 24 pacientes dos conjuntos A (saúáveis, olhos abertos) e C (doentes, olhos abertos). A partir disso, aplicamos o mapeamento no estudo para as séries relacionadas as ondas teta, beta, alfa e delta. Analogamente às séries originais, como cada série possui $T = 1.024$ pontos, usamos $Q = 20$ quantis e $k = 1, 2, \dots, 25$ nas novas simulações. Para as séries (teta, beta, alfa e delta) e para os conjuntos A e C, calculamos os valores médios de $C(k)$, $M(k)$, $\Delta(k)$ e $B(k)$ versus k e seus respectivos desvios padrões, sendo esses representados por barras de erros para os pacientes saúáveis e os pacientes doentes (Figs. 23, 24, 25 e 26). Note que, a curva associada à pacientes saúáveis (A) e com a DA (C) formam dois clusters distintos com máxima separação para um dado valor de k e para a onda cerebral delta (Figs. 23 (d), 24 (d), 25 (d) e 26 (d)). É conhecido na literatura que os estágios iniciais da DA são caracterizados pelo aumento da atividade das ondas teta e pela diminuição da atividade das ondas beta. E também que, à medida que a doença progride, tais mudanças são acompanhadas pela diminuição das ondas alfa. Um aumento das ondas delta ocorre no estágio mais avançado da DA. Portanto, concluímos que a base de dados usada neste trabalho é composta por pacientes que apresentam características de um estágio mais avançado da doença e que o mapeamento em estudo foi capaz de diferenciar as características do estágio da DA.

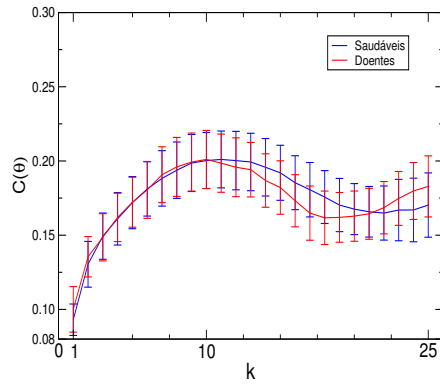
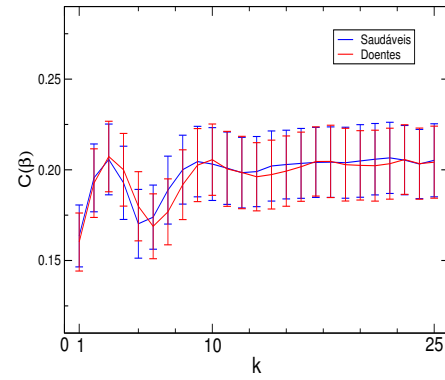
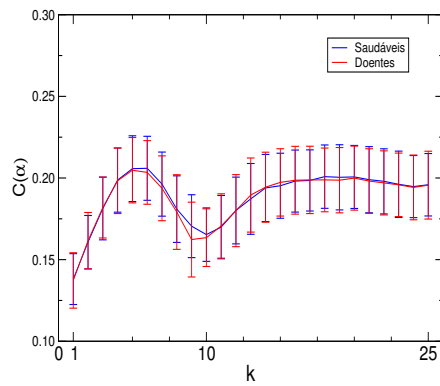
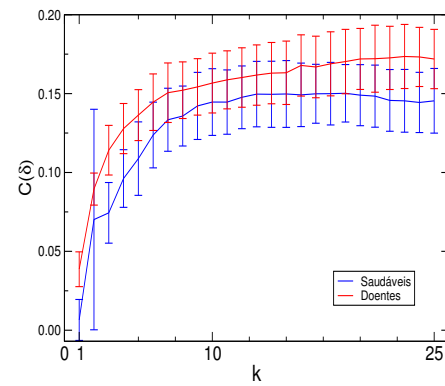
(a) *Teta*(b) *Beta*(c) *Alfa*(d) *Delta*

Figura 23: $C(\theta)$, $C(\beta)$, $C(\alpha)$ e $C(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos).

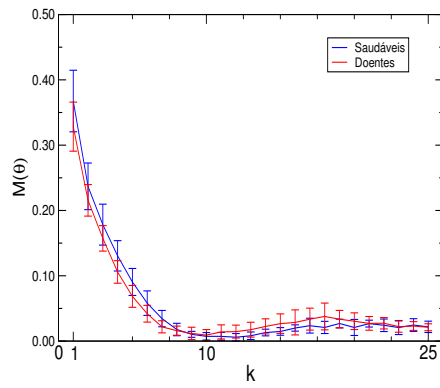
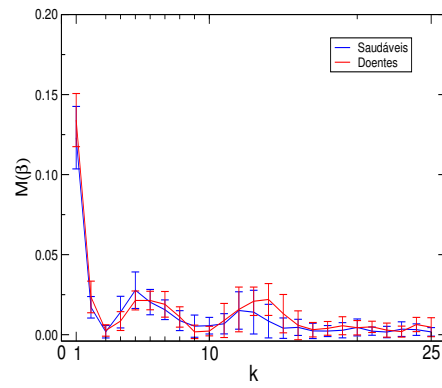
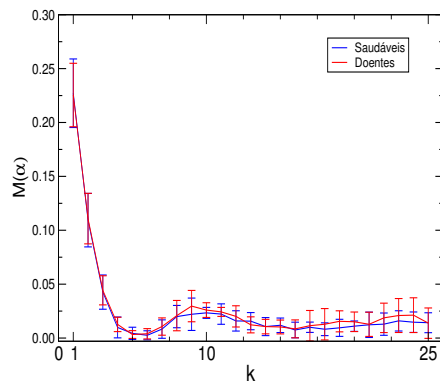
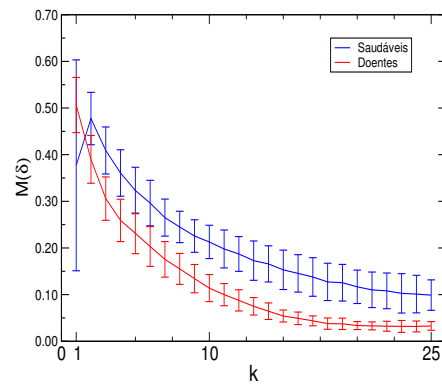
(a) *Teta*(b) *Beta*(c) *Alfa*(d) *Delta*

Figura 24: $M(\theta)$, $M(\beta)$, $M(\alpha)$ e $M(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos).

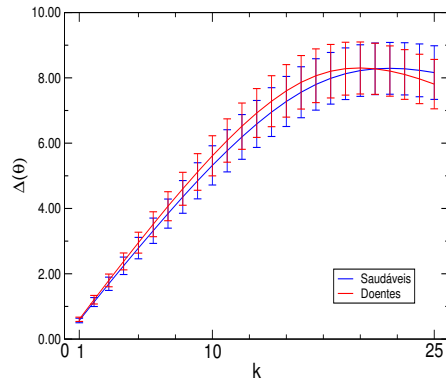
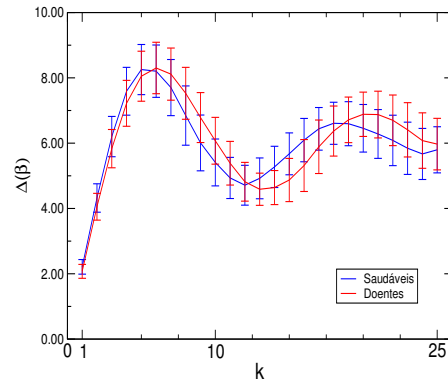
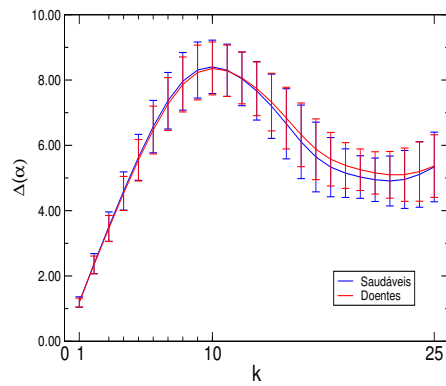
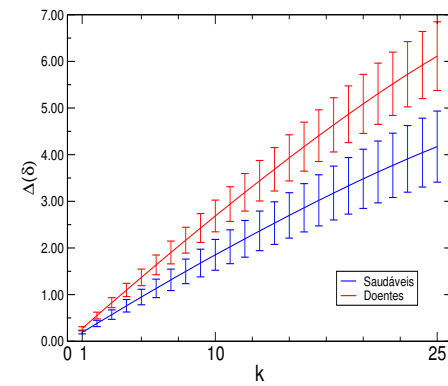
(a) *Teta*(b) *Beta*(c) *Alfa*(d) *Delta*

Figura 25: $\Delta(\theta)$, $\Delta(\beta)$, $\Delta(\alpha)$ e $\Delta(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos).

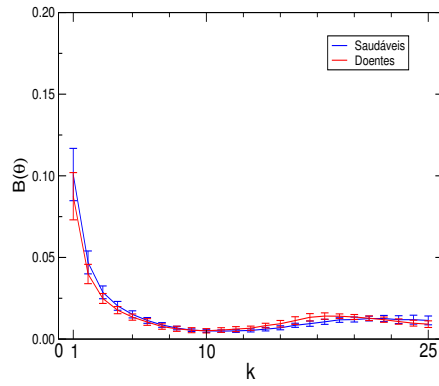
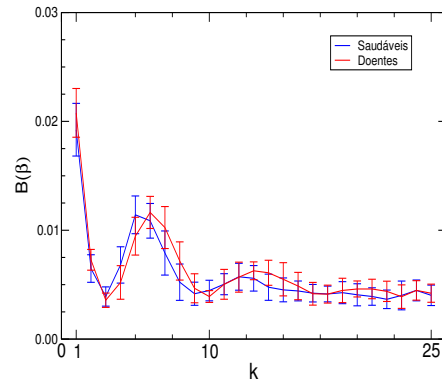
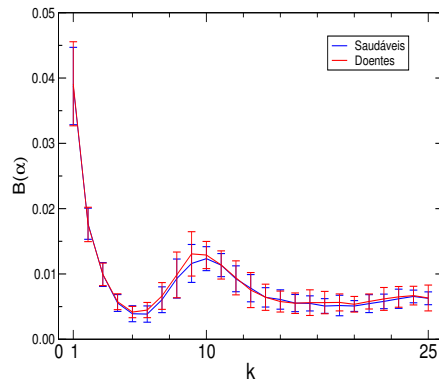
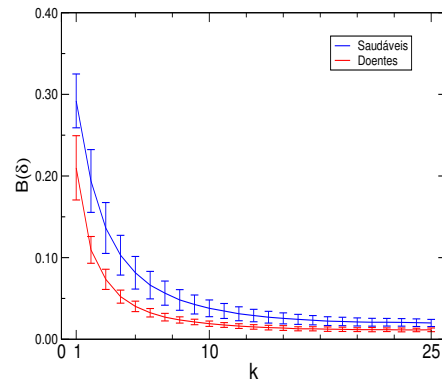
(a) *Teta*(b) *Beta*(c) *Alfa*(d) *Delta*

Figura 26: $B(\theta)$, $B(\beta)$, $B(\alpha)$ e $B(\delta)$ versus k , (canal P_3) com $T = 1.024$, $Q = 20$ e $k = 1, 2, \dots, 25$ para os conjuntos A (pacientes saudáveis, olhos abertos) e C (pacientes doentes, olhos abertos).

4.3 Discussão

Com base nos resultados apresentados, o mapeamento em estudo foi capaz de diferenciar envelhecimento da doença de Alzheimer, independentemente do caracterizador e do eletrodo utilizados. Nesse sentido, o comprimento do salto médio que possui complexidade linear e que, conseqüentemente, demanda menor

esforço computacional, se mostrou o caracterizador com o melhor desempenho em tal diferenciação.

No tocante à investigação dos possíveis estágios da DA, embora os resultados apresentados nos levem a conjecturar que todos os indivíduos com a DA possuem a doença em seu estágio mais avançado, a ausência de informações clínicas de cada paciente, tais como as idades e as avaliações cognitivas, nos impede de confirmar essa conjectura.

5 CONCLUSÕES

Uma técnica de diagnóstico robusta e não-invasiva da DA, com base na análise de dados de EEG, foi proposta neste trabalho. Os resultados obtidos com o uso dessa técnica são estatisticamente significantes e atestam que a metodologia empregada pode ser uma ferramenta útil na análise de sinais complexos e não-lineares, como os produzidos pelos EEGs de humanos. De modo geral, as análises realizadas mostraram que o mapeamento em estudo foi capaz de distinguir com êxito pacientes em diferentes condições patológicas. Mais especificamente, os quatro caracterizadores utilizados, a saber, C , M , Δ e B mostraram que esse mapeamento foi capaz de distinguir pacientes saudáveis (com olhos abertos ou fechados) de pacientes com a DA (com olhos abertos ou fechados) sob uma ampla variação dos valores de k e para todos os canais de EEG da base de dados utilizada.

Os caracterizadores topológicos em estudo para as redes associadas aos pacientes no estágio mais avançado da DA (ondas delta) apresentaram valores topológicos distintos dos das redes associadas aos pacientes saudáveis. Para as demais ondas, tais valores foram próximos. Isso nos levou a conjecturar que todos os indivíduos com a DA possuem a doença em seu estágio mais avançado. Contudo, a ausência de informações clínicas de cada paciente, tais como as idades e as avaliações cognitivas, nos impede de confirmar essa conjectura.

Em trabalhos futuros, desejamos utilizar o mapeamento no estudo da identificação dos primeiros sinais da DA. Nesse sentido, esperamos que os caracterizadores topológicos em estudo associados aos pacientes com a DA em seu estágio inicial apresentem valores distintos dos obtidos para os pacientes saudáveis e pacientes nos estágios mais avançados da DA.

REFERÊNCIAS BIBLIOGRÁFICAS

World Health Organization. <https://www.who.int/>, 2012. [Online; acesso em 14 de fevereiro de 2017].

ACHARYA, R.; FAUST, O.; KANNATHAL, N.; CHUA, T.; LAXMINARAYAN, S. Non-linear analysis of EEG signals at various sleep stages. **Computer methods and programs in biomedicine**, v.80, n.1, p.37–45, 2005.

ADELI, H.; GHOSH-DASTIDAR; DADMEHR, N. Alzheimer's Disease: Models of Computation and Analysis of EEGs. **Clinical EEG and NeuroScience**, v.36, 2005.

AHMADLOU, M.; ADELI, H.; ADELI, A. New diagnostic EEG markers of the Alzheimer's Disease using visibility graph. **Journal of Neural Transmission**, v.117, 2010.

AL-FAHOUM, A. S.; AL-FRAIHAT, A. A. Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. **ISRN neuroscience**, v.2014, 2014.

AMARAL, L. A. N.; IVANOV, P. C.; AOYAGI, N.; HIDAKA, I.; TOMONO, S.; GOLDBERGER, A. L.; STANLEY, H. E.; YAMAMOTO, Y. Behavioral-independent features of complex heartbeat dynamics. **Physical Review Letters**, v.86, 2001.

AMARAL, L. A. N.; SOARES, D. J. B.; DA SILVA, L. R.; LUCENA, L. S.; SAITO, M.; KUMANO, H.; AOYAGI, N.; YAMAMOTO, Y. Power law temporal auto-

correlations in day-long records of human physical activity and their alteration with disease. **Europhys. Lett.**, v.66, 2004.

Anaconda, 2011. Anaconda Navigator. <https://www.anaconda.com/>, 2011. [Online; acesso em 9 de abril de 2018].

BAGNATO, G. G. Análise estrutural de Redes Complexas Modulares por Meio de Caminhadas Auto-Excludentes. Universidade de São Paulo - São Carlos, 2018. Tese de doutorado - Universidade de São Paulo.

BRANDES, U. A faster algorithm for betweenness centrality. **Journal of mathematical sociology**, v.25, n.2, p.163–177, 2001.

BUDSON, A.; SOLOMON, P. **Memory Loss, Alzheimer’s Disease, and Dementia**. New York: Elsevier, 2015.

CAMPANHARO, A. S.; DOESCHER, E.; RAMOS, F. M. Application of Quantile Graphs to the Automated Analysis of EEG Signals. **Neural Processing Letters**, 2018.

CAMPANHARO, A. S. L. O. Dualidade Entre a Análise de Séries Temporais e de Redes Complexas. INPE - São José dos Campos, 2011. Tese de doutorado - Instituto Nacional de Pesquisas Espaciais.

CAMPANHARO, A. S. L. O.; DOESCHER, E.; RAMOS, F. M. Automated EEG Signals Analysis Using Quantile Graphs. In: ROJAS, I.; JOYA, G.; CATALA, A., (Ed.), **Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science**, Springer International Publishing, 10306v., 2017.

CAMPANHARO, A. S. L. O.; RAMOS, F. M. Hurst exponent estimation of self-affine time series using quantile graphs. **Physica A**, v.444, 2016a.

CAMPANHARO, A. S. L. O.; RAMOS, F. M. Quantile graphs for the characterization of chaotic dynamics in time series. In: **WCCS 2015 — IEEE Third World Conference on Complex Systems**, IEEE, 2016b.

CAMPANHARO, A. S. L. O.; SIRER, M. I.; MALMGREN, R. D.; RAMOS, F. M.; AMARAL, L. A. N. Duality between Time Series and Networks. **PLoS ONE**, v.6, 2011.

CANNADY, J. Artificial neural networks for misuse detection. In: **National information systems security conference**, Baltimore, 26v., 1998.

CASTANHO, M.; BARROS, L.; VENDITE, L.; YAMAKAMI, A. Avaliação de um teste em medicina usando uma curva ROC fuzzy. **Biomatemática**, v.14, p.19–28, 2004.

Code::Blocks, 2008. The IDE with all the features you need, having a consistent look, feel and operation across platforms. <http://www.codeblocks.org/>, 2008. [Online; acesso em 15 de outubro de 2017].

COSTA, L. F.; RODRIGUES, F. A.; TRAVIESO, G.; VILLAS, P. R. Characterization of complex networks. **Advances in Physics**, v.56, 2007.

DAUBECHIES, I. The wavelet transform, time-frequency localization and signal analysis. **IEEE transactions on information theory**, v.36, n.5, p.961–1005, 1990.

DURONGBHAN, P.; ZHAO, Y.; CHEN, L.; ZIS, P.; DE MARCO, M.; UNWIN, Z. C.; VENNERI, A.; HE, X.; LI, S.; ZHAO, Y.; ET AL. A Dementia Classification Framework using Frequency and Time-frequency Features based on EEG signals. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, 2019.

ECKMANN, J. P.; RUELLE, D. Fundamental limitations for estimating dimensions and Lyapunov exponents in dynamical systems. **Physica D**, v.56, p.185–187, 1992.

FAGIOLO, G. Clustering in complex directed networks. **Physical Review E**, v.76, 2007.

FELDMAN, H. H.; WOODWARD, M. The staging and assessment of moderate to severe Alzheimer disease. **Neurology**, v.65, 2005.

GAO, Z.; JIN, N. Complex network from time series based on phase space reconstruction. **Chaos**, v.19, 2009.

GOLUB, G. H.; VAN LOAN, C. F. Matrix computations. **The Johns Hopkins University Press, Baltimore, USA**, 1989.

GUIMERÁ, R.; SALES-PARDO, M.; AMARAL, L. Modularity from fluctuations in random graphs and complex networks. **Physical Review E**, v.70, 2004.

GUTIERREZ, B. A. O.; SILVA, H. S.; GUIMARÃES, C.; CAMPINO, A. C. Economic impact of Alzheimer's Disease in Brazil: is it possible to improve care and minimize costs? **Ciência & Saúde Coletiva**, v.19, 2014.

HIGOR B. SANTOS, M. V. S.; STEDILE, R. Sistema de Identificação de Padrões de Ondas Cerebrais. Departamento Acadêmico de Eletrônica, 2014. Trabalho de conclusão de curso - Universidade Tecnológica Federal do Paraná.

HOUMANI, N.; VIALATTE, F.; GALLEGUO-JUTGLÀ, E.; DREYFUS, G.; NGUYEN-MICHEL, V.-H.; MARIANI, J.; KINUGAWA, K. Diagnosis of Alzheimer's disease with Electroencephalography in a differential framework. **PloS one**, v.13, n.3, 2018.

HUSAIN, M.; GARRETT, R. Clinical Diagnosis and Management of Alzheimer's Disease. **Neuroimaging Clinics of North America**, v.15, 2005.

IVANOV, P. C.; AMARAL, L. A. N.; GOLDBERGER, A. L.; HAVLIN, S.; ROSENBLUM, M. G.; STANLEY, H. E.; STRUZIK, Z. R. From $1/f$ noise to multifractal cascades in heartbeat dynamics. **Chaos**, v.11, 2000.

IVANOV, P. C.; AMARAL, L. A. N.; GOLDBERGER, A. L.; STANLEY, H. E. Stochastic feedback and the regulation of biological rhythms. **Europhysics Letters**, v.43, 1998.

KANTZ, H.; SCHREIBER, T. **Nonlinear time series analysis**. Cambridge: Cambridge University Press, 2003.

KORNER, T. W. **Fourier analysis**. Cambridge: Cambridge University Press, 1988.

LACASA, L.; LUQUE, B.; BALLESTEROS, F.; LUQUE, J.; NUÑO, J. C. From time series to complex Networks. **Proceedings of the National Academy of Sciences**, v.105, 2008.

LAI, C.; CHUNG, P.; TSENG, V. S. A novel two-level clustering method for time series data analysis. **Expert Systems with Applications**, v.37, 2010.

LO, C.-C.; AMARAL, L. A. N.; HAVLIN, S.; IVANOV, P. C.; PENZEL, T.; PETER, J.-H.; STANLEY, H. E. Dynamics of sleep-wake transitions during sleep. **Europhysics Letters**, v.57, p.625–631, 2002.

NEWMAN, M. E.; GIRVAN, M. Finding and evaluating community structure in networks. **Physical review E**, v.69, n.2, p.026113, 2004.

NEWMAN, M. E. J. **Networks: An Introduction**. New York: Oxford University Press, 2010.

PADOVANI, C. R. **Delineamento de Experimentos**. São Paulo: Cultura Acadêmica, 2014.

PERCIVAL, D. B.; WALDEN, A. T. **Wavelet methods for time series analysis**. Cambridge: Cambridge University Press, 2000.

PINEDA, A. M.; RAMOS, F. M.; BETTING, L. E.; CAMPANHARO, A. S. Use of Complex Networks for the Automatic Detection and the Diagnosis of Alzheimer's Disease. In: Springer, 2019. **International Work-Conference on Artificial Neural Networks**; resumos. , 2019. 115–126.

PRATI, R.; BATISTA, G.; MONARD, M. Curvas ROC para avaliação de classificadores. **Revista IEEE América Latina**, v.6, n.2, p.215–222, 2008.

PRITCHARD, W. S.; DUKE, D. W.; COBURN, K. L.; MOORE, N. C.; TUCKER, K. A.; JANN, M. W.; HOSTETLER, R. M. EEG-based, neural-net predictive classification of Alzheimer's disease versus control subjects is augmented by non-linear EEG measures. **Electroencephalography and clinical Neurophysiology**, v.91, n.2, 1994.

RODRIGUES, F. A. Caracterização, classificação e análise de redes complexas. São Carlos, 2007. Tese (Doutorado) - Universidade de São Paulo, Instituto de Física de São Carlos.

RStudio, 2017. RStudio. <https://www.rstudio.com/>, 2017. [Online; acesso em 19 de julho de 2019].

SANTOS-MAYO, L.; SAN-JOSÉ-REVUELTA, L. M.; ARRIBAS, J. I. A computer-aided diagnosis system with EEG based on the P3b wave during an auditory odd-ball task in schizophrenia. **IEEE transactions on biomedical engineering**, v.64, n.2, p.395–407, 2016.

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**, v.27, n.2, 1996.

STROGATZ, S. H. **Nonlinear dynamics and chaos**. Boulder: Westview Press, 1994.

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**, v.2014, 2014.

TU, J. V. Advantages and disadvantages of using artificial neural networks versus logistic regression for predicting medical outcomes. **Journal of clinical epidemiology**, v.49, n.11, 1996.

TURNER, P. Grace. <http://plasma-gate.weizmann.ac.il/Grace/>, 1995. [Online; acesso em 15 de outubro de 2017].

VERPLANCKE, T.; LOOY, S. V.; STEURBAUT, K.; BENOIT, D.; TURCK, F. D.; MOOR, G. D.; DECRUYENAERE, J. A novel time series analysis approach for prediction of dialysis in critically ill patients using echo-state networks. **BMC Medical Informatics and Decision Making**, v.10, 2010.

XU., X.; ZHANG, J.; SMALL, M. Superfamily phenomena and motifs of networks induced from time series. **Proceedings of the National Academy of Sciences of the United States of America**, v.105, 2008.

YAGNESWARAN, S.; BAKER, M.; PETROSIAN, A. Power frequency and wavelet characteristics in differentiating between normal and Alzheimer EEG. In: **Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference 2002. Proceedings of the Second Joint, IEEE**, 2002.

ZHANG, J.; LUO, X.; SMALL, M. Detecting chaos in pseudoperiodic time series without embedding. **Physical Review E**, v.73, 2006.

ZHANG, J.; SMALL, M. Complex Network from Pseudoperiodic Time Series. **Physical Review Letters**, v.96, 2006.

Apêndice A - Testes Estatísticos

ANOVA

A análise de variância (ANOVA) tem como objetivo comparar as médias das populações amostrais, e assim identificar se essas médias diferem significativamente. Enquanto que nos testes de hipótese se trabalha com duas amostras, a ANOVA compara as médias de mais de duas amostras e determina se ao menos uma se difere significativamente das demais. Assim, o teste ANOVA permite que vários grupos sejam comparados a um só tempo. Este teste é uma extensão do teste t de Student (que compara apenas duas médias) para um número qualquer de médias. A estatística do teste para a ANOVA é calculada por meio do teste F (Fisher-Snedecor). Algumas pressuposições básicas precisam ser satisfeitas para o uso desse teste, ou seja, os erros são variáveis aleatórias independentes, a variância é constante (homogênea nos tratamentos) e a distribuição dos erros é normal ou aproximadamente normal (Padovani, 2014).

ÁREA SOB A CURVA ROC

Uma ferramenta amplamente utilizada para avaliar o desempenho de testes de diagnósticos é a análise da curva ROC (Receiver Operating Characteristic) (Castanho et al., 2004). Quando consideramos o resultado de um teste de diagnóstico para duas populações, uma população com a doença e outra sem, dificilmente observamos uma perfeita separação entre esses grupos. Portanto, qualquer que seja o “valor de corte” (ponto que separa as duas populações) escolhido, alguns indivíduos com a doença serão classificados corretamente como verdadeiros positivos

(VP) e alguns serão classificados como falsos negativos (FN). Por outro lado, alguns casos sem a doença serão classificados como falsos positivos (FP) mas alguns serão corretamente classificados como verdadeiros negativos (VN).

As medidas de sensibilidade e especificidade fornecem a probabilidade de que o teste classifique corretamente um indivíduo doente e um indivíduo saudável, respectivamente. Tais medidas são dadas por:

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (17)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (18)$$

Supondo que os resultados de testes médicos sejam traduzidos por um valor em um determinado intervalo contínuo, o ponto de corte que separa os diagnósticos verdadeiros e falsos deve pertencer a este intervalo. A área sob a curva ROC é, então empregada como uma medida do desempenho do teste como discriminador de pacientes doentes e saudáveis. Um teste ideal é aquele cuja área sob a curva ROC é igual a 1. Quando a curva ROC é a bissetriz, ou seja, área igual a 0,5, o teste não permite distinguir os grupos (Prati et al., 2008).

Os testes estatísticas utilizados neste trabalho foram executadas no ambiente de programação e na linguagem R (RStudio, 2017).