



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM EM CIÊNCIAS
DA TERRA E DO AMBIENTE
Doutorado em Modelagem em Ciências da Terra e do Ambiente

Tese de doutorado

ρ_{DCCA} e DMC_x^2 : Implementação, Otimização e Aplicações

Apresentada por: Fernando Ferraz Ribeiro
Orientador: Prof. Dr. Gilney Figueira Zebende
Co-orientador: Prof. Dr. Juan Alberto Leyva Cruz

Maio de 2025

Fernando Ferraz Ribeiro

ρ_{DCCA} e DMC_x^2 : Implementação, Otimização e Aplicações

Tese de doutorado apresentada ao Programa de Pós-graduação em Modelagem em Ciências da Terra e do Ambiente, Curso de Doutorado em Modelagem em Ciências da Terra e do Ambiente da UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA, como requisito parcial para a obtenção do título de **Doutor em Modelagem em Ciências Ambientais**.

Área de conhecimento: Estudos Ambientais e Geotecnologias

Orientador: Prof. Dr. Gilney Figueira Zebende

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Co-orientador: Prof. Dr. Juan Alberto Leyva Cruz

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Programa de Pós-graduação em Modelagem em Ciências da Terra e do Ambiente

Doutorado em Modelagem em Ciências da Terra e do Ambiente

A Banca Examinadora, constituída pelos professores listados abaixo, leu e recomenda a aprovação da Tese de doutorado, intitulada “ \boldsymbol{P}_{DCCA} e DMC_x^2 : Implementação, Otimização e Aplicações”, apresentada no dia 23 de maio de 2025, como requisito parcial para a obtenção do título de **Doutor em Modelagem em Ciências Ambientais**.

Orientador:

Prof. Dr. Gilney Figueira Zebende

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Co-Orientador:

Prof. Dr. Juan Alberto Leyva Cruz

UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Membro externo da Banca:

Prof. Dr. Thadeu Josino Pereira Penna

UNIVERSIDADE FEDERAL FLUMINENSE

Membro externo da Banca:

Prof. Dr. Jês de Jesus Fias Cerqueira

UNIVERSIDADE FEDERAL DA BAHIA

Membro externo da Banca:

Prof. Dr. Alexsandro Guerra Cerqueira

UNIVERSIDADE FEDERAL DA BAHIA

Membro externo da Banca:

Prof. Dr. —

UNIVERSIDADE —

Dedico este trabalho a...

Agradecimentos

Ao Prof. Dr. Gilney Figueira Zebende , pela orientação e dedicação ao tema escolhido, e por ter acreditado na pesquisa .

Ao Prof. Dr. Juan Alberto Leyva Cruz

A UEFS e ao PPGM pelos recursos proporcionados para a elaboração da pesquisa

Feira de Santana, BA, Brasil
23 de Maio de 2025

Fernando Ferraz Ribeiro

Resumo

Escreva aqui o seu resumo em português.

Palavras Chaves: Séries Temporais, Clima , DMC_x^2 , $\boldsymbol{\rho}_{DCCA}$, Ciência de Dados, Aprendizado de Máquina

Abstract

Write here your abstract in english.

Keywords: Time Series, Climate, DMC_x^2 , $\boldsymbol{\rho}_{DCCA}$, Data Science, Machine Learning

Sumário

1	Introdução	1
1.1	Definição do problema	2
1.2	Objetivos	2
1.3	Importância da Pesquisa	2
1.4	Viabilidade e Limitações	3
1.5	Questões e Hipóteses	4
1.6	Metodologia	5
1.7	Organização da Tese	5
2	Fundamentação Teórica e Metodologia	7
2.1	Métodos de Análise de Séries Temporais: DFA , $DCCA$, ρ_{DCCA} e DMC_x^2	7
2.2	Fractais e os métodos baseados no DFA	10
2.3	Variantes de correlação cruzada baseados no $DCCA$	12
2.4	Aplicações dos métodos	14
3	Artigo 01:	15
3.1	Artigos	15
4	Artigo 02:	16
5	Artigo 03:	17
6	Conclusão	18
A	Anexo I	19
A.1	Produção técnica	19
	Referências	21

Lista de Tabelas

Lista de Figuras

2.1	Análise <i>DFA</i> de sequências de DNA e sequências de controle, resultados apresentando correlação de longo alcance (+ e \circ) e não apresentando correlação. (\times e \square)	12
-----	--	----

Lista de Quadros

Lista de Algoritmos

Lista de Siglas

UEFS	Universidade Estadual de Feira de Santana
PPGM	Programa de Pós-Graduação em Modelagem em Ciências da Terra e do Ambiente
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
<i>DCCA</i>	<i>Detrended Cross-Correlation Analysis</i>
<i>DFA</i>	<i>Detrended Fluctuation Analysis</i>
DMC_x^2	<i>Detrended Multiple Cross-Correlation Coefficient</i>
ρ_{DCCA}	<i>Detrended Cross-Correlation Coefficient</i>
DPDCCA ...	<i>Detrended partial Cross-Correlation Analysis</i>

Introdução

“Ordinary life is pretty complex stuff.”
(Harvey Pekar)

Os sistemas complexos compreendem um campo interdisciplinar da ciência, cujo escopo de problemas que trata não possui uma definição exata. Este conjunto amplo de fenômenos é comumente identificado e agrupado por algumas de suas características: são formados pela contribuição de um conjunto (geralmente grande) de componentes (muitas vezes simples) que, interagindo, estruturam-se de forma auto-organizada, gerando resultados inesperados, que não podem ser previstos pelos estudos estatísticos e/ou matemáticos tradicionais dos elementos formadores do sistema.

Na área dos estudos ambientais, os sistemas complexos aparecem em problemas de diversas naturezas: sistemas de transportes, redes de energia e comunicação, organizações sociais e econômicas, densidade e ocupação humana do espaço, dentre outras. Os estudos do clima e de variáveis meteorológicas ocupam um espaço de particular relevância na intercessão entre os estudos ambientais e os sistemas complexos. Em 2021, a Academia Real das Ciências da Suécia concedeu metade do Prêmio Nobel de Física para Syukuro Manabe e Klaus Hasselmann, cujos estudos apresentam modelos complexos para a análise do clima, em particular apontando para uma correlação entre as emissões de dióxido de carbono e as mudanças climáticas.

Muitos fenômenos complexos são investigados pela análise de grandes conjuntos de dados. É notável a velocidade e quantidade de dados que são gerados, obtidos e armazenados pela humanidade atualmente. A aquisição, manipulação, gestão, armazenamento e criação de valor a partir de dados, através de ambientes computacionais, tem-se apresentado como um novo paradigma tecnológico. Um campo do conhecimento que recebeu a denominação de Ciência de Dados, conceito que envelopa alguns termos frequentemente associados à inovação científica, técnica e social como *Big Data*, mineração de dados, *Business Intelligence* internet das coisas, inteligência artificial (IA) e aprendizado de máquina (AM), dentre outros ([EMC EDUCATION SERVICE, 2015](#), p. 12-13).

Uma das formas em que os dados costumam ser organizados são as denominadas séries temporais. As séries temporais são definidas como um conjunto de observações (numéricas ou categóricas) ordenado no tempo. Embora muitos dos dados que descrevem as dinâmicas espaciais podem ser registrados na forma de séries temporais (abastecimento de água nas tubulações, consumo de energia elétrica nos imóveis, fluxos de pessoas e veículos pela ci-

dade, casos de uma doença por dia, etc.), contudo as técnicas de medição de correlações, bem como a devida exploração destas para inferir novos conhecimentos, permanecem como perguntas abertas em muitas sub-áreas das ciências ambientais (Bermudez-Edo; BARNAGHI; MOESSNER, 2018).

1.1 Definição do problema

As séries temporais registram aspectos fundamentais de uma série de problemas, de diversas áreas, que, devidamente abordados, podem clarificar características e comportamentos destas estruturas de dados. O coeficiente DMC_x^2 é uma ferramenta promissora para a análise de correlações em séries temporais múltiplas, permitindo a identificação de padrões complexos que não são capturados por métodos tradicionais. No entanto, sua aplicação em problemas reais ainda é limitada devido à falta de implementações computacionais eficientes e estudos que explorem seu potencial em diferentes contextos.

1.2 Objetivos

Neste trabalho, tem-se por objetivo investigar a aplicabilidade do ρ_{DCCA} e do DMC_x^2 , com o objetivo, não só de identificar correlações entre variáveis, mas também de entender o algoritmo, propor e aperfeiçoar ferramentas computacionais para a realização dos cálculos.

Como objetivos gerais foram elencados:

1. Testar os algoritmos e os métodos;
2. implementar um algoritmo computacional geral para calcular o DMC_x^2 para qualquer número de séries temporais;
3. testar as ferramentas implementadas em novas análises.

1.3 Importância da Pesquisa

A pesquisa que propõe melhorias no algoritmo e a implementação de uma ferramenta para o cálculo do $DCCA$, ρ_{DCCA} e DMC_x^2 é de grande relevância, pois aborda um problema computacional desafiador e ainda pouco explorado. A eficiência no cálculo destes métodos é essencial para viabilizar sua aplicação em conjuntos de dados de grande escala, que são cada vez mais comuns em diversas áreas do conhecimento. Além disso, a disponibilização

de uma ferramenta computacional robusta e acessível pode democratizar o uso do método, permitindo que pesquisadores de diferentes áreas possam utilizá-lo sem a necessidade de desenvolver suas próprias implementações.

A proposta de melhorias no algoritmo também tem implicações diretas na precisão e na escalabilidade das análises realizadas com o DMC_x^2 . Métodos computacionalmente custosos podem limitar o tamanho dos conjuntos de dados analisados ou mesmo inviabilizar estudos em contextos onde o tempo de processamento é crítico. Ao otimizar o cálculo do DMC_x^2 , a pesquisa contribui para superar essas limitações, abrindo caminho para novas aplicações e para a exploração de padrões complexos em séries temporais múltiplas.

No contexto das ciências ambientais, as implicações são particularmente significativas. A análise de séries temporais de variáveis climáticas e meteorológicas é fundamental para entender fenômenos como mudanças climáticas, padrões de precipitação e distribuição de energia solar. As redes elétrica, de água e de dados são fontes de informações únicas sobre a dinâmica de uma cidade. A mobilidade de pessoas e mercadorias são problemas desafiadores para gestores, planejadores e empresas privadas. Séries temporais também são exploradas no contexto de dados sísmicos, tanto para antevisão e planejamento diante de uma catástrofe natural, quanto para o mapeamento do sub-solo e identificação de recursos geológicos. Uma ferramenta eficiente para o cálculo do DMC_x^2 pode auxiliar na identificação de correlações entre variáveis ambientais ou com forte implicação no espaço e nas dinâmicas locais, contribuindo para o desenvolvimento de modelos analíticos mais precisos e para a formulação de políticas públicas baseadas em dados confiáveis.

1.4 Viabilidade e Limitações

A implementação e melhoria dos algoritmos propostos apresentam alta viabilidade, uma vez que não demandam recursos computacionais excepcionais. A maior parte das operações envolvidas, como manipulação de matrizes e cálculos estatísticos, pode ser realizada em computadores domésticos de configurações medianas, resultando em um baixo investimento financeiro.

A viabilidade como análise de hipóteses existe. Certamente limitações como o peso computacional são mais difíceis de antever, mas devem ser levadas em conta desde o princípio.

A obtenção de conjuntos de dados reais e confiáveis é um desafio significativo que pode limitar o progresso de pesquisas como esta. A disponibilidade de dados de qualidade frequentemente depende de fatores externos, como acesso a fontes restritas, custos associados à aquisição ou limitações técnicas na coleta e armazenamento.

De resto, vale lembrar que a honestidade do trabalho científico pode levar a resultados que validam, total ou parcialmente, ou invalidam as conjecturas iniciais. Apenas com a atenta avaliação dos experimentos realizados pode-se entender os ganhos obtidos no percurso. A análise dos dados meteorológicos será aplicada para descrever esse percurso, independente do status da validação dos resultados.

1.5 Questões e Hipóteses

Este trabalho foi baseado em algumas premissas:

1. O DMC_x^2 , pelas características de análise do método, pode ajudar a entender características de séries temporais e aplicado em problemas de diversas áreas.
2. O DMC_x^2 é uma generalização do método ρ_{DCCA} para múltiplas séries temporais.
3. O ρ_{DCCA} , em determinadas condições testadas, apresentou resultados mais interessantes (como melhor descrição dos fenômenos) que os apresentados pelo coeficiente de Pearson quando aplicado à séries temporais (WANG et al., 2013).

Partindo destas premissas, procuramos responder perguntas basilares:

1. Poderia-se criar uma ferramenta computacional eficaz para o cálculo dos coeficientes?
2. Poderia-se encontrar uma alternativa mais eficaz o cálculo dos coeficientes, principalmente o recente DMC_x^2 ?
3. A implementação e melhorias, se bem sucedidas, poderiam ampliar o número de possíveis investigações em que os coeficientes podem ser aplicados?

Para orientar o trabalho, duas hipóteses foram formuladas:

- 1.
2. Uma ferramenta adequada para o cálculo dos coeficientes seria um impulso na aplicação, utilização e divulgação dos métodos estudados.

1.6 Metodologia

A metodologia utilizada neste projeto insere-se dentro da abordagem metodológica experimental, estruturada em três etapas principais.

A primeira etapa consiste em uma revisão bibliográfica abrangente, com o objetivo de compreender os fundamentos teóricos e metodológicos relacionados ao DFA , $DCCA$, ρ_{DCCA} e DMC_x^2 . Nesta fase, busca-se consolidar o entendimento sobre os algoritmos e metodologias vigentes, bem como identificar lacunas e oportunidades de melhoria nos métodos existentes.

Na segunda etapa, realiza-se uma avaliação detalhada das possibilidades e limitações das ferramentas computacionais disponíveis para o cálculo dos coeficientes. Essa avaliação se dará através de uma pesquisa que aplica os métodos, utilizando as ferramentas de cálculo existente. Este procedimento deve embasar o pesquisador para analisar, criticar e avaliar as possibilidades e limitações das ferramentas, das pesquisas e das perguntas que podem ser formuladas e respondidas através do ferramental vigente.

Com base nessa análise, é proposta a implementação de uma nova ferramenta computacional, com foco na otimização do algoritmo, na praticidade do uso e na eficiência do cálculo. Essa etapa envolve o desenvolvimento, teste e validação da ferramenta, garantindo que ela seja robusta e capaz de lidar com conjuntos de dados de diferentes tamanhos e complexidades.

Por fim, na terceira etapa, a ferramenta desenvolvida é aplicada a um conjunto de dados complexo, com o objetivo de testar os avanços proporcionados pela implementação. Essa aplicação prática permite avaliar a eficácia das melhorias realizadas, bem como explorar novas possibilidades de análise e interpretação de padrões em séries temporais múltiplas.

Essa abordagem metodológica experimental permite não apenas validar as hipóteses formuladas, mas também contribuir para o avanço do estado da arte na análise de correlações em séries temporais, com implicações práticas em diversas áreas do conhecimento.

1.7 Organização da Tese

Esta tese foi formulada no formato *multi-paper*. Após uma breve introdução do objeto de estudo, realizada no presente capítulo, este trabalho apresenta:

Uma revisão da literatura, no Capítulo 2, abrangendo nos métodos de correlação entre

séries temporais basados no *DFA*, com um foco específico no estudo de métodos que propõe alguma forma de correlação entre as séries.

Entre os Capítulos 3 e 5, uma sequência de pesquisas que compõem o cerne da tese. São trabalhos independentes porem interconectados que são apresentados em seguir.

O Capítulo 3 apresenta um artigo publicado, apresentando uma aplicação do ρ_{DCCA} e do DMC_x^2 na análise de múltiplas séries temporais. Uma contribuição para a análise de séries de ondas cerebrais, captadas por eletroencefalograma (EEG) também é apresentada.

O Capítulo 4 apresenta um artigo ainda não publicado apresenta a criação de uma ferramenta computacional que possibilita velocidade e *usabilidade* no cálculo dos algoritmos do *DFA*, *DCCA* e ρ_{DCCA} . Além da implementação da biblioteca, o artigo apresenta como contribuição, uma mudança na sequência de passos extremamente vantajosa para o cálculo do *DCCA* e ρ_{DCCA} , principalmente para cálculo com muitas séries paralelas, como no caso da atribuição de valores na matriz do ρ_{DCCA} para cálculo do DMC_x^2 .

O Capítulo 5...

O Capítulo 6 por fim, resume os achados e analisa os resultados.

Fundamentação Teórica e Metodologia

“Time is nature’s way of keeping everything
from happening at once.”
(John Archibald Wheeler)

O presente capítulo apresenta uma revisão bibliográfica dos métodos utilizados, iniciando por uma explicação passo a passo do método para calcular o DFA e o $DCCA$, sem os quais o ρ_{DCCA} e o DMC_x^2 não podem ser entendidos, seguindo pela explanação do cálculo do ρ_{DCCA} e do DMC_x^2 , como se dá a leitura dos resultados e como são aplicados em análises.

Em seguida apresenta-se as relações dos métodos com a geometria fractal e multifractal, conceitos importantes nas análises das séries temporais propostas pelos métodos derivados do DFA . Apresentamos também metodologia e algoritmo para da determinação da multifractalidade de uma série temporal, segundo o DFA .

Na Sessão 2.3, apresenta-se uma seleção de outros algoritmos que, além do DMC_x^2 , trabalham com correlações múltiplas baseadas no DFA . Algoritmos selecionados são apresentados.

Na Sessão 2.4, apresenta-se aplicações do que está família de metodologias já foi aplicada, ressaltando a gama de problemas e os mais relacionados cm as ciências ambientais.

2.1 Métodos de Análise de Séries Temporais: DFA , $DCCA$, ρ_{DCCA} e DMC_x^2

O coeficiente ρ_{DCCA} (ZEBENDE, 2011) foi formulado tendo como bases o *Detrended Fluctuation Analysis* (DFA) (PENG et al., 1994a) e o *Detrended Cross-Correlation Analysis* ($DCCA$) (PODOBNIK; STANLEY, 2008). O DFA é um método de análise de uma série temporal que fornece um parâmetro de auto-afinidade. O termo *Detrended* refere-se a eliminação de uma tendência. O processo é executado em 6 passos:

1. **Cálculo da série integrada:** dada uma série temporal $\{x_i\}$, com i variando de 1 a N , a série integrada X_k é calculada por $X_k = \sum_{i=1}^k [x_i - \langle x \rangle]$, com k também

variando de 1 a N ;

2. **Divisão da série em caixas:** a série integrada X_k é dividida em $N - n$ caixas de tamanho n (escala temporal), cada caixa contendo $n + 1$ valores, começando em i até $i + n$;¹
3. **Cálculo do ajuste polinomial:** para cada caixa, é calculado polinômio (geralmente de grau 1) que melhor se ajusta, obtendo $\tilde{X}_{k,i}$ com $i \leq k \leq (i + n)$;
4. **Cálculo da função f_{DFA}^2 para cada caixa:** para cada caixa de uma escala temporal é calculada a função de DFA pela expressão:

$$f_{DFA}^2(n, i) = \frac{1}{1+n} \sum_{k=i}^{i+n} (X_k - \tilde{X}_k)^2;$$

5. **Cálculo do DFA para uma escala temporal:** para todas as caixas de uma escala de tempo, o DFA é calculado como:

$$F_{DFA}(n) = \sqrt{\frac{1}{N-n} \sum_{i=1}^{N-n} f_{DFA}^2(n, i)};$$

6. **Análise em diferentes escalas temporais:** para diferentes escalas de tempo (n), com valores possíveis $4 \leq n \leq \frac{N}{4}$, é calculada a função F_{DFA} para encontrar uma relação entre $F_{DFA} \times n$

O DFA também representa as propriedades de auto-correlação de longo alcance de uma lei de potência (ZEBENDE; SILVA; FILHO, 2013). Existe uma proporção entre o DFA e o expoente α dada por $F_{DFA}(n) \propto n^\alpha$, portando uma proporção entre $\log(F_{DFA}(n)) \propto \alpha$. Diferentes valores de α podem ser interpretados como:

- $\alpha < 1/2$: Anti-correlação,
- $\alpha \simeq 1/2$: Sem correlação(white noise),
- $\alpha > 1/2$: Correlação,
- $\alpha \simeq 1$: ruído rosa
- $\alpha > 1$: Série não estacionária,
- $\alpha \simeq 3/2$: ruído browniano.

O ajuste do polinômio, definido como a tendência, costuma utilizar um polinômio de primeira ordem, mas existem que comparam o efeito que a ordem do polinômio exerce sobre a análise e a capacidade do polinômio de maior grau capturar a tendência da série (HU et al., 2001).

¹Aqui considera-se o cálculo com sobreposição de caixas.

O $DCCA$ amplia o DFA para estabelecer a correlação entre duas séries temporais (PODOBNIK; STANLEY, 2008). O valor deste coeficiente tende a ser a média dos valores do DFA das duas séries e segue os 6 passos descritos abaixo:

1. **Cálculo das séries integradas:** tomando duas séries temporais com a mesma extensão $\{x\alpha_i\}$ e $\{x\beta_i\}$ com i variando de 1 a N , as séries integradas $X\alpha_k$ e $X\beta_k$ são calculadas por $X_k = \sum_{i=1}^k [x_i - \langle x \rangle]$ para cada série, com k também variando de i a N ;
2. **Divisão das séries em caixas:** as séries $X\alpha_k$ e $X\beta_k$ são divididas em $N - n$ caixas de tamanho n (escala de tempo), cada caixa contendo $n + 1$ valores, começando em i até $i + n$;
3. **Cálculo dos ajustes de polinômios:** para cada caixa, um polinômio (geralmente de grau 1) melhor se ajusta, obtendo $\widetilde{X\alpha_{k,i}}$ e $\widetilde{X\beta_{k,i}}$, para séries $\{x\alpha_i\}$ e $\{x\beta_i\}$ respectivamente, com $i \leq k \leq (i + n)$;
4. **Cálculo da função f_{DCCA}^2 em cada caixa:** para cada caixa uma das $N - n$ caixas de uma mesma escala temporal a função é calculada por:
$$f_{DCCA}^2(n, i) = \frac{1}{1+n} \sum_{k=i}^{i+n} (X\alpha_k - \widetilde{X\alpha_{k,i}}) \times (X\beta_k - \widetilde{X\beta_{k,i}})$$
5. **Cálculo do $DCCA$ para toda a escala temporal:** para todas as caixas de uma mesma escala temporal, o $DCCA$ é calculado como:
$$F_{DCCA}(n) = \sqrt{\frac{1}{N-n} \sum_{i=1}^{N-n} f_{DCCA}^2(n, i)};$$
6. **Análise em diferentes escalas temporais:** para um número de escalas de tempo (n), com valores possíveis $4 \leq n \leq \frac{N}{4}$, o $DCCA$ é calculado para encontrar uma relação entre $F_{DCCA} \times n$

Este coeficiente λ indica a existência de uma correlação entre duas séries regidas por leis de potência, mas não quantifica o nível desta correlação. O *Detrended cross-correlation coefficient* ou ρ_{DCCA} (equação 2.1) é um coeficiente que, variando entre -1 e 1, aponta ausência de correlação cruzada para valores próximos de zero, sendo maior a correlação quanto mais o valor se aproximar de 1 e maior a anti-correlação quanto mais o valor se aproximar de -1 (ZEBENDE, 2011).

$$\rho_{DCCA}(n) = \frac{F_{DCCA}^2(x\alpha, x\beta)(n)}{F_{DFA}(x\alpha)(n) \times F_{DFA}(x\beta)(n)} \quad (2.1)$$

O método foi estatisticamente validado (PODOBNIK et al., 2011), testado (VASSOLER; ZEBENDE, 2012; GUEDES et al., 2017; FERREIRA et al., 2018), e critérios para ava-

liação de relevância estatísticas do resultados foram desenvolvidos (GUEDES et al., 2018a; GUEDES et al., 2018b).

O ρ_{DCCA} foi estendido para calcular a correlação cruzada de múltiplas series temporais. Denominado *Detrended Multiple Cross-Correlation Coefficient* (DMC_x^2), representa a generalização do ρ_{DCCA} para múltiplas variáveis (ZEBENDE; SILVA, 2018). Também foi implementado com abordagem de janelas móveis (GUEDES; da Silva Filho; ZEBENDE, 2021) e foi desenvolvido um teste estatístico para o coeficiente múltiplo (da Silva Filho et al., 2021)

A generalização utiliza a equação 2.2. Sendo y uma serie temporal definida como variável dependente, x_j um conjunto de séries temporais tomadas como variáveis independentes, com j variando de 1 a m . Sendo $\rho_{y,x_j}(n)$ o vetor coluna contendo dos valores dos coeficientes ρ_{DCCA} entre a serie y e cada uma das séries x_j (Eq: 2.3).

$$DMC_x^2 \equiv \rho_{Y,X_j}(n)^T \times \rho^{-1}(n) \times \rho_{Y,X_j}(n) \quad (2.2)$$

$$\rho_{Y,X_j}(n)^T = [\rho_{Y,X_1}(n), \rho_{Y,X_2}(n), \dots, \rho_{Y,X_m}(n)] \quad (2.3)$$

A matriz $\rho^{-1}(n)$, inversa da matriz apresentada na equação 2.4. A matriz é montada através dos valores dos coeficientes ρ_{DCCA} para cada par de variáveis dependentes em x_w . Pela propriedade comutativa presente na equação do cálculo do ρ_{DCCA} , pode-se afirmar que a matriz é sempre simétrica. O valor do ρ_{DCCA} de uma série com ela mesma é sempre o de correlação total (1), portando a diagonal principal apresenta o valor 1 em todas as posições.

$$\rho(n) = \begin{pmatrix} 1 & \rho_{X_1,X_2}(n) & \rho_{X_1,X_3}(n) & \dots & \rho_{X_1,X_m}(n) \\ \rho_{X_2,X_1}(n) & 1 & \rho_{X_2,X_3}(n) & \dots & \rho_{X_2,X_m}(n) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ \rho_{X_m,X_1}(n) & \rho_{X_m,X_2}(n) & \rho_{X_m,X_3}(n) & \dots & 1 \end{pmatrix} \quad (2.4)$$

2.2 Fractais e os métodos baseados no DFA

Existe uma relação entre os métodos baseados no DFA e a geometria fractal, termo cunhado pelo matemático Benoit Mandelbrot em um artigo publicado em 1975. A palavra fractal procede do latim *fractus* cujo significado é "quebrado" ou "fragmentado".

A pesquisa que originou a geometria fractal baseou-se em pesquisas anteriores do matemático Gaston Julia, e aplicou computação gráfica para estudar o comportamento de métodos iterativos na resolução de certas equações complexas. O conjunto de imagens geradas pelo experimento apresentou características e propriedades geométricas particulares, dentre elas a persistência do valor obtido pelo cálculo da dimensão de Hausdorff (que, a partir dos trabalhos de Mandelbrot, passa a ser conhecida também como dimensão fractal), ainda que calculado utilizando diferentes escalas de discretização do espaço de resultados (MANDELBROT, 1982).

A técnica de contagem de caixas, também conhecida como método de contagem de boxes, é uma ferramenta utilizada para estimar a dimensão fractal de uma forma. Técnica foi introduzida por Mandelbrot e tem sido amplamente utilizada em estudos de fractais.

A ideia básica da contagem de caixas é dividir o conjunto em caixas de tamanho fixo e contar o número de caixas que o conjunto ocupa. Em seguida, o tamanho da caixa é reduzido e o processo é repetido. A relação entre o número de caixas e o tamanho da caixa pode ser usada para estimar a dimensão fractal do conjunto.

A fórmula para a dimensão fractal utilizando a técnica de contagem de caixas é dada por:

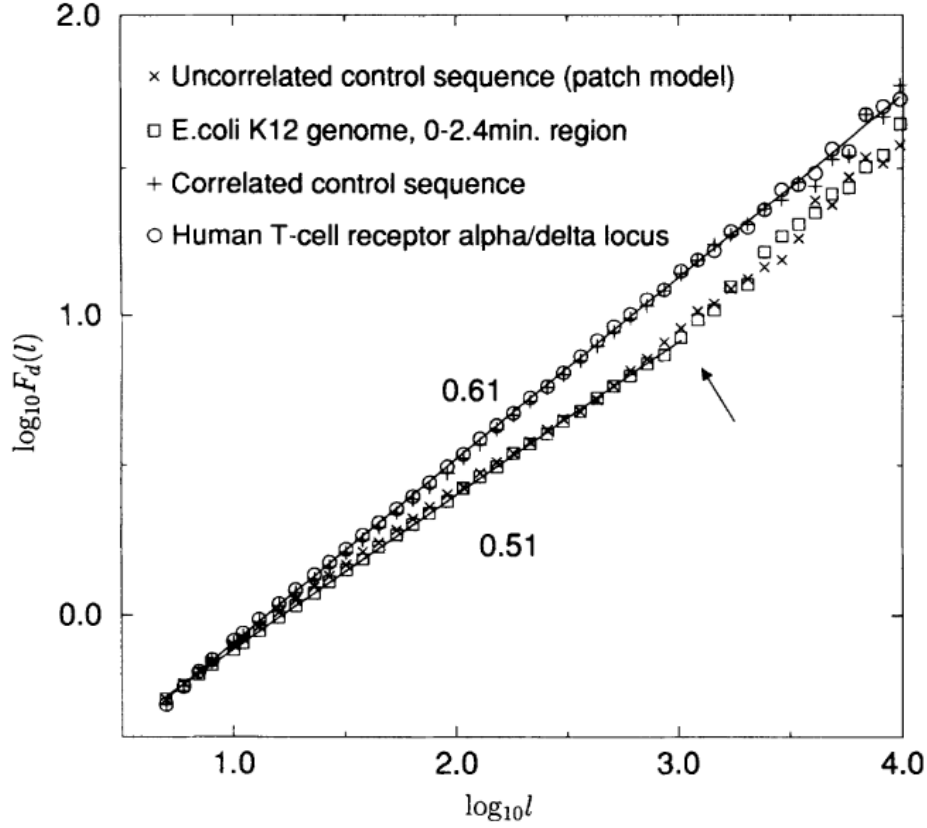
$$D = \lim_{\epsilon \rightarrow 0} \frac{\log(N(\epsilon))}{\log(1/\epsilon)} \quad (2.5)$$

Sendo $N(\epsilon)$ é o número de caixas de tamanho ϵ que preenchem o conjunto.

A ideia de se dividir uma figura 2D em caixas apresenta semelhança com a divisão de uma série temporal (1D) por segmentos de reta de um polinômio de primeiro grau. Mais do que uma semelhança aleatória, em escala log-log, o *DFA*, quando desenha um segmento de reta ligando uma sequência de escalas temporais, representa uma característica de auto-similaridade da série, como mostrado na Figura 2.1. ()

Uma generalização multifractal do *DFA* foi proposta por Kantelhardt et al. (2002). Apresentado na Equação 2.6, onde o valor de q representa a ordem da função *DFA* calculada. Quando $q = 2$ a equação assume o valor do *DFA*. No caso de $q = 0$ a média exponencial é utilizada para a medição. Esses resultados também foram expandidos, baseando-se no *DCCA* para a correlação multifractal entre séries temporais (ZHOU, 2008).

Figura 2.1: Análise *DFA* de seqüências de DNA e seqüências de controle, resultados apresentando correlação de longo alcance (+ e \circ) e não apresentando correlação. (\times e \square)



Fonte: (PENG et al., 1994b)

$$q \neq 0,$$

$$F_q(s) \equiv \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(v, s)]^{q/2} \right\}^{1/q},$$

(2.6)

$$q = 0,$$

$$F_0(s) \equiv \exp \left\{ \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln [F^2(v, s)] \right\} \sim s^{h(0)}$$

2.3 Variantes de correlação cruzada baseados no *DCCA*

Duas variantes do *DCCA* que lidam com correlação cruzada já foram apresentados nesta pesquisa: O ρ_{DCCA} e o DMC_x^2 . Além desses métodos, este trabalho destaca mais alguns trabalhos: *Multivariate Detrended Fluctuation Analysis (MVDFA)* (XIONG; SHANG,

2017), *Generic Multichannel DFA* (*GM DFA*) (NAVEED; REHMAN, 2025) e o *Detrended Partial Cross-correlation Coefficient* (*DPCCA*) (YUAN et al., 2015).

Diferente do DMC_x^2 , que estabelece um valor de função, onde a correlação é medida entre um conjunto de variáveis independentes e uma variável dependente, o *MVDFA* calcula um valor único para qualquer combinação de séries temporais, representando uma generalização do *DFA* para múltiplas variáveis. Neste método, divide-se a série temporal em um conjunto de caixas não sobreposto, eventualmente perdendo alguns valores no final da série, mas repetindo o processo do fim para o começo da reta para compensar. Após a interpolação dos polinômios, a função local, em cada caixa e dada pela Equação 2.7, onde a norma euclidiana (Equação 2.8) é utilizada para o cálculo.

$$F^2(v, s) = \frac{1}{s} \sum_{t=1}^s \left\| (Y_{l_v+t,1}, Y_{l_v+t,2}, \dots, Y_{l_v+t,p}) - (\tilde{Y}_{.,1}, \tilde{Y}_{.,2}, \dots, \tilde{Y}_{.,p}) \right\|^2 \quad (2.7)$$

$$\|X - Y\| = \|(x_1, x_2, \dots, x_n) - (y_1, y_2, \dots, y_n)\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (2.8)$$

Terminando o cálculo de forma semelhante ao apresentado no algoritmo do *DFA*, como mostrado na Equação 2.9. O *MG DFA* substitui a norma euclideana pela norma de Mahalanobis. Estes métodos não possuem qualquer relação com o *DCCA*, por extensão, com o ρ_{DCCA} .

$$F_{MVDFA}(s) = \left\{ \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F^2(v, s)] \right\}^{1/n} \quad (2.9)$$

Já o *DPDCCA* aplica a mesma matriz utilizada no cálculo do DMC_x^2 , apresentado na Equação 2.2, cuja forma antes da inversão é apresentada na Equação 2.4. A matriz é denominada de $C(n)$ (Equação 2.10).

$$C(n) = \rho^{-1}(n) = \begin{pmatrix} C_{1,1}(n) & C_{1,2}(n) & \dots & C_{1,m}(n) \\ C_{2,1}(n) & C_{2,2}(n) & \dots & C_{2,m}(n) \\ \vdots & \vdots & & \vdots \\ C_{m,1}(n) & C_{m,2}(n) & \dots & C_{m,m}(n) \end{pmatrix} \quad (2.10)$$

O cálculo final do *DPDCCA* é realizado pela Equação 2.11, a partir dos valores obtidos

pela inversão da matriz do $\boldsymbol{\rho}_{DCCA}$.

$$\rho_{DPCCA}(j_1, j_2; n) = \frac{-C_{j_1, j_2}(n)}{\sqrt{C_{j_1, j_1}(n) \cdot C_{j_2, j_2}(n)}} \quad (2.11)$$

O resultado deve ser analisado como a correlação entre as duas variáveis, desconsiderando a atuação das outras variáveis cujos valores de $\boldsymbol{\rho}_{DCCA}$ foram utilizados na montagem da matriz na Equação 2.4.

2.4 Aplicações dos métodos

Artigo 01:

“That brain of mine is something more than merely mortal,
as time will show.”
(Ada Lovelace)

3.1 *Artigos*

Artigo 02:

“Just remember, once you’re over the hill
you begin to pick up speed.”
(Charles M. Schulz)

Artigo 03:

“Noel Meyerhof consulted the list he had prepared
and chose which item was to be first.
As usual, he relied mainly on intuition.”
(Isaac Asimov)

Conclusão

“I live on Earth at present, and I don’t know what I am.

I know that I am not a category.

I am not a thing – a noun.

I seem to be a verb, an evolutionary process -

an integral function of the universe.”

(R. Buckminster Fuller)

Anexo I

Aqui pode ser um Título do seu Anexo

Texto do Anexo.

A.1 Produção técnica



REPÚBLICA FEDERATIVA DO BRASIL
MINISTÉRIO DO DESENVOLVIMENTO, INDÚSTRIA, COMÉRCIO E SERVIÇOS
INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL
DIRETORIA DE PATENTES, PROGRAMAS DE COMPUTADOR E TOPOGRAFIAS DE CIRCUITOS

Certificado de Registro de Programa de Computador

Processo Nº: **BR512023001376-5**

O Instituto Nacional da Propriedade Industrial expede o presente certificado de registro de programa de computador, válido por 50 anos a partir de 1º de janeiro subsequente à data de 15/05/2023, em conformidade com o §2º, art. 2º da Lei 9.609, de 19 de Fevereiro de 1998.

Título: Multiple EEG cross-correlation

Data de publicação: 15/05/2023

Data de criação: 01/02/2022

Titular(es): GILNEY FIGUEIRA ZEBENDE; ANDREA DE ALMEIDA BRITO; FLORÊNCIO MENDES OLIVEIRA FILHO; FERNANDO FERRAZ RIBEIRO; JUAN ALBERTO LEYVA CRUZ

Autor(es): GILNEY FIGUEIRA ZEBENDE; ANDREA DE ALMEIDA BRITO; FLORÊNCIO MENDES OLIVEIRA FILHO; JUAN ALBERTO LEYVA CRUZ; FERNANDO FERRAZ RIBEIRO

Linguagem: PYTHON

Campo de aplicação: MT-06

Tipo de programa: SM-01

Algoritmo hash: SHA-256

Resumo digital hash: 28a265f9b7e4b4ff4e40d57f8fb671cb2dd941d7c6579e607b4392d362b64cbd

Expedido em: 23/05/2023

Aprovado por:

Carlos Alexandre Fernandes Silva
Chefe da DIPTO

Referências Bibliográficas

- Bermudez-Edo, M.; BARNAGHI, P.; MOESSNER, K. Analysing real world data streams with spatio-temporal correlations: Entropy vs. Pearson correlation. *Automation in Construction*, v. 88, n. May 2017, p. 87–100, 2018. ISSN 09265805. [1](#)
- da Silva Filho, A. M. et al. Statistical test for Multiple Detrended Cross-Correlation Coefficient. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 562, p. 125285, 2021. ISSN 03784371. [2.1](#)
- EMC EDUCATION SERVICE. *Data Science and Big Data Analytics: Discovering, Analyzing, Visualizing and Presenting Data*. Indianapolis, Indiana, USA: JOHN WILEY & SONS, 2015. ISBN 978-1-118-87605-3 1-118-87605-9 978-1-118-87613-8 1-118-87613-X 978-1-118-87622-0 1-118-87622-9. [1](#)
- FERREIRA, P. et al. A sliding windows approach to analyse the evolution of bank shares in the European Union. *Physica A: Statistical Mechanics and its Applications*, v. 490, p. 1355–1367, 2018. ISSN 03784371. [2.1](#)
- GUEDES, E. et al. DCCA cross-correlation in blue-chips companies: A view of the 2008 financial crisis in the Eurozone. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 479, p. 38–47, 2017. ISSN 03784371. [2.1](#)
- GUEDES, E. F. et al. Statistical test for $\Delta\rho$ DCCA cross-correlation coefficient. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 501, p. 134–140, 2018. ISSN 03784371. [2.1](#)
- GUEDES, E. F. et al. Statistical test for $\Delta\rho$ DCCA: Methods and data. *Data in Brief*, v. 18, p. 795–798, 2018. ISSN 23523409. [2.1](#)
- GUEDES, E. F.; da Silva Filho, A. M.; ZEBENDE, G. F. Detrended multiple cross-correlation coefficient with sliding windows approach. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 574, p. 125990, 2021. ISSN 03784371. [2.1](#)
- HU, K. et al. Effect of trends on detrended fluctuation analysis. *Physical Review E*, v. 64, n. 1, p. 011114, jun. 2001. ISSN 1063-651X, 1095-3787. [2.1](#)
- KANTELHARDT, J. W. et al. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A*, 2002. [2.2](#)
- MANDELBROT, B. B. *The Fractal Geometry of Nature*. 1982. [S.l.]: Henry Holt and Company, 1982. ISBN 978-0-7167-1186-5. [2.2](#)
- NAVEED, K.; REHMAN, N. U. Fully multivariate detrended fluctuation analysis using Mahalanobis norm with application to multivariate signal denoising. *Measurement*, v. 242, p. 116142, jan. 2025. ISSN 02632241. [2.3](#)
- PENG, C.-K. et al. Mosaic Organization of DNA Nucleotides. v. 49, n. 2, p. 1685–1689, 1994. [2.1](#)

- PENG, C.-K. et al. Mosaic organization of DNA nucleotides. *Physical Review E*, v. 49, n. 2, p. 1685–1689, fev. 1994. ISSN 1063-651X, 1095-3787. [2.1](#)
- PODOBNIK, B. et al. Statistical tests for power-law cross-correlated processes. *Phys. Rev. E*, American Physical Society, v. 84, n. 6, p. 66118, 2011. [2.1](#)
- PODOBNIK, B.; STANLEY, H. E. Detrended cross-correlation analysis: A new method for analyzing two nonstationary time series. *Physical Review Letters*, v. 100, n. 8, 2008. ISSN 00319007. [2.1](#), [2.1](#)
- VASSOLER, R. T.; ZEBENDE, G. F. DCCA cross-correlation coefficient apply in time series of air temperature and air relative humidity. *Physica A*, Elsevier B.V., v. 391, n. 7, p. 2438–2443, 2012. ISSN 0378-4371. [2.1](#)
- WANG, G. J. et al. Random matrix theory analysis of cross-correlations in the US stock market: Evidence from Pearson’s correlation coefficient and detrended cross-correlation coefficient. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 392, n. 17, p. 3715–3730, 2013. ISSN 03784371. [3](#)
- XIONG, H.; SHANG, P. Detrended fluctuation analysis of multivariate time series. *Communications in Nonlinear Science and Numerical Simulation*, v. 42, p. 12–21, jan. 2017. ISSN 10075704. [2.3](#)
- YUAN, N. et al. Detrended Partial-Cross-Correlation Analysis: A New Method for Analyzing Correlations in Complex System. *Scientific Reports*, v. 5, n. 1, p. 8143, jan. 2015. ISSN 2045-2322. [2.3](#)
- ZEBENDE, G. F. DCCA cross-correlation coefficient: Quantifying level of cross-correlation. *Physica A: Statistical Mechanics and its Applications*, Elsevier B.V., v. 390, n. 4, p. 614–618, 2011. ISSN 03784371. [2.1](#), [2.1](#)
- ZEBENDE, G. F.; SILVA, A. M. Detrended Multiple Cross-Correlation Coefficient. *Physica A*, Elsevier B.V., v. 510, p. 91–97, 2018. ISSN 0378-4371. [2.1](#)
- ZEBENDE, G. F.; SILVA, M. F.; FILHO, A. M. DCCA cross-correlation coefficient differentiation : Theoretical and practical approaches. *Physica A*, v. 392, p. 1756–1761, 2013. [2.1](#)
- ZHOU, W.-X. Multifractal detrended cross-correlation analysis for two nonstationary signals. *Physical Review E*, v. 77, n. 6, p. 066211, jun. 2008. ISSN 1539-3755, 1550-2376. [2.2](#)

ρ_{DCCA} e DMC_x^2 : Implementação, Otimização e Aplicações

Fernando Ferraz Ribeiro

Feira de Santana, BA, Maio de 2025.