UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

Decomposição de séries temporais preservando o viés determinístico

Felipe Simões Lage Gomes Duarte

Tese de Doutorado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-CCMC)

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Felipe Simões Lage Gomes Duarte

Decomposição de séries temporais preservando o viés determinístico

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Rodrigo Fernandes de Mello

Coorientador: Prof. Dr. Eduardo Raul Hruschka

USP – São Carlos
Janeiro de 2020

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

D812d Duarte, Felipe Simões Lage Gomes
Decomposição de séries temporais preservando o
viés determinístico / Felipe Simões Lage Gomes
Duarte; orientador Rodrigo Fernandes de Mello;
coorientador Eduardo Raul Hruschka. -- São Carlos,
2019.
120 p.

Tese (Doutorado - Programa de Pós-Graduação em
Ciências de Computação e Matemática Computacional) --
Instituto de Ciências Matemáticas e de Computação,
Universidade de São Paulo, 2019.

1. Séries temporais. 2. Decomposição de séries
temporais. 3. Análise de séries temporais. 4.
Predição de séries temporais. I. Mello, Rodrigo
Fernandes de, orient. II. Hruschka, Eduardo Raul,
coorient. III. Título.

Felipe Simões Lage Gomes Duarte

Time series decomposition while preserving deterministic influences

Doctoral dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Doctorate Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Rodrigo Fernandes de Mello

Co-advisor: Prof. Dr. Eduardo Raul Hruschka

USP – São Carlos
January 2020

À minha esposa Karina, familiares e amigos.

AGRADECIMENTOS

À minha esposa Karina, que sempre me apoiou e me deu forças para superar todos os obstáculos desta jornada (e que jornada). Ela esteve ao meu lado nos momentos fáceis e batalhou comigo nos mais difíceis. A ela reconheço grande parte desta conquista. Te amo!

Aos meus pais, Rafael e Edna, que em nenhum momento mediram esforços para realização dos meus sonhos. Guiaram-me competentemente pelos caminhos corretos da vida, ensinando a fazer as melhores escolhas e mostrando-me que a honestidade e o respeito são virtudes essenciais à vida. A eles devo a pessoa que me tornei, sou extremamente feliz e tenho muito orgulho por chamá-los de pai e mãe. Amo vocês!

Ao meu irmão Bernardo Duarte que sempre me ajudou e esteve ao meu lado no que foi preciso. Apesar de vivermos grande parte das nossas vidas em cidades diferentes, nos divertimos a cada encontro, sorrimos em cada ligação e brindamos em cada vitória. Não poderia ter irmão melhor!

Ao meu sogro Edson, minha sogra Elisabete (*in memoriam*), meu cunhado Edson Jr. e sua esposa Thaís Cunha por todo carinho e atenção dedicados a mim e acolhendo-me como integrante de sua família.

Aos amigos e parceiros que a vida me deu: Marco, Leila e a turminha, Diego e Karen, Diandra, Adelmo, Tieta e Edvaldo, Tiago e Pina, Fernando (Fêu), Fábio Sikansi, Francisco (Chicão), Samuel (Fadel), Tácito (Paréa) e Niedja, Duzinho e Lessandra, demais amigos do Itaú Unibanco e Santander Brasil.

Ao meu orientador e grande amigo Prof. Dr. Rodrigo Fernandes de Mello pela paciência, orientação e risadas nas reuniões. Ele que fez o possível (e às vezes o impossível) para tornar o sonho da pós-graduação realidade para mim e é um exemplo a ser seguido. Obrigado!

Aos professores e amigos Dr. Eduardo Raul Hruschka e Dr. Ricardo Araújo Rios pelas diversas discussões teóricas, incansáveis revisões e reuniões nos horários mais importunos.

À FAPESP (Processo 2014/21636-3 e 2013/07375-0) e Capes pelo aporte financeiro. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade do(s) autor(es) e não necessariamente refletem a visão da FAPESP e da CAPES.

*“Por vezes sentimos que aquilo que fazemos não é senão uma gota de água no mar.
Mas o mar seria menor se lhe faltasse uma gota.”
(Madre Teresa de Calcuta)*

RESUMO

DUARTE, FELIPE S. L. G. **Decomposição de séries temporais preservando o viés determinístico**. 2020. 120 p. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Avanços tecnológicos possibilitaram a coleta e modelagem de grandes quantidades de dados produzidos ao longo do tempo por fenômenos industriais, humanos e naturais. Em se tratando de séries temporais, tais dados são compostos por influências determinísticas, relacionadas a eventos recorrentes e unicamente dependentes de observações passadas, e estocásticas, associadas a efeitos aleatórios. Modelos produzidos com base em apenas uma dessas influências tendem a produzir resultados sub-ótimos e incompletos. Portanto, idealmente, deve-se modelar o componente estocástico por meio de ferramentas estatísticas e o determinístico utilizando ferramentas da área de Sistemas Dinâmicos. Esse cenário leva à inerente necessidade da decomposição de dados temporais, em busca de modelos mais acurados e melhores resultados de predição. Diversas abordagens têm sido utilizadas para realizar tal decomposição, tais como: (i) Transformada de Fourier; (ii) Transformadas Wavelet; (iii) Médias Móveis; (iv) Análise Espectral Singular; (v) Lazy; (vi) GHKSS; e (vii) outras abordagens baseadas no método de decomposição de modo empírico (EMD – *Empirical Mode Decomposition*). Tais abordagens apresentam problemas associados à imposição de viés definido pelos seus conjuntos de funções admissíveis, sendo que o senoidal é predominante sobre o componente determinístico resultante, descaracterizando o viés original dos dados e levando a modelagens sub-ótimas, consequentemente gerando resultados insatisfatórios para o processo de predição. Neste contexto, esta tese de doutorado introduz três abordagens de decomposição de séries temporais que visam preservar, ao máximo, as influências determinísticas por meio da utilização de espaços-fase, resultando em representações mais fiéis do viés original dos dados: (i) **Spring**, (ii) **Spring Time Domain** e (iii) **Spring***. Essas abordagens foram experimentalmente avaliadas e comparadas ao estado da arte com base em métricas comumente adotadas na literatura, mais precisamente: Média do Erro Absoluto (do inglês *Mean Absolute Error* – MAE) e Distância Média da Linha Diagonal (do inglês *Mean Distance from Diagonal Line* – MDDL). Spring e suas variantes comprovaram ser mais eficazes para a segmentação entre influências determinísticas e estocásticas, naturalmente levando à melhoria do processo de modelagem e predição de séries temporais. Por fim, para validar a hipótese de que as decomposições propostas melhoram resultados de predição, as abordagens foram conectadas às técnicas de modelagem polinomial e de funções de base radial, permitindo reduzir significativamente erros decorrentes do processo de previsão.

Palavras-chave: Séries temporais, Decomposição de séries temporais, Análise de séries temporais, Predição de séries temporais.

ABSTRACT

DUARTE, FELIPE S. L. G. **Time series decomposition while preserving deterministic influences**. 2020. 120 p. Tese (Doutorado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2020.

Technological advances allowed to collect and model large scales of data produced along time by industrial, human and natural phenomena. In terms of time series, such data are composed of deterministic, solely based on past observations or associated to recurrent events, and stochastic influences, due to random effects. Models using only one of those influences tend to produce sub-optimal and incomplete results. Therefore, one should ideally model the stochastic component using Statistical tools and the deterministic one with Dynamical Systems. This scenario leads us to the decomposition of time series in attempt to obtain more accurate models and best prediction results. Several approaches have been applied to address such decomposition stage, including: (i) Fourier Transform; (ii) Wavelet Transform; (iii) Moving Average; (iv) Singular Spectrum Analysis; (v) Lazy; (vi) GHKSS; and (vii) other approaches based on the Empirical Mode Decomposition (EMD). Those approaches have drawbacks related to the bias imposed by their respective sets of admissible functions, having the sinusoidal as the typical to represent the deterministic component extracted, thus loosing the original time series bias, what leads to sub-optimal models and insatisfactory prediction results. In this context, this PhD thesis introduces three time series decompositions approaches that aim to preserve as much as possible the deterministic influences by using phase-spaces, what helps maintaining the original data bias, to mention: (i) **Spring**, (ii) **Spring Time Domain** e (iii) **Spring***. Those approaches were experimentally assessed and compared against the state-of-the-art through measurements commonly used in the literature, more precisely: Mean Absolute Error (MAE) and Mean Distance from Diagonal Line (MDDL). Spring and its variations confirmed to be more effective to separate deterministic and stochastic influences, thus improving the modeling and prediction processes. At last, the proposed decomposition approaches were plugged into the polynomial and the radial basis function prediction techniques to confirm the hypothesis that the forecasting of series observations could be more accurate what was corroborated given errors were overall reduced.

Keywords: Time series, Time series decomposition, Time series analysis, Time series prediction.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de série temporal produzida por componentes determinísticos e estocásticos.	27
Figura 2 – Componentes produzidos pela técnica EMD-RP.	28
Figura 3 – Viés senoidal imposto pelas atuais técnicas de decomposição.	28
Figura 4 – Diferença série temporal contínua e discreta.	34
Figura 5 – Diferença de séries estacionária, não-estacionária com tendência linear e não-estacionária com tendência exponencial.	36
Figura 6 – Exemplo da sensibilidade ao estado inicial em séries caóticas.	41
Figura 7 – Diferença no atrator final provocado pela diferença do estado inicial em séries caóticas.	42
Figura 8 – Estimação da dimensão de separação.	45
Figura 9 – Efeitos sobre o atrator para a escolha da dimensão de separação.	46
Figura 10 – Atrator de Rössler.	48
Figura 11 – Gráfico de percentual de falsos vizinhos.	48
Figura 12 – Processo de decomposição utilizando a Transformada de Fourier.	53
Figura 13 – Função <i>wavelet</i> formada pela derivada de segunda ordem da função gaussiana.	55
Figura 14 – Decomposição da série temporal utilizando médias móveis.	57
Figura 15 – Etapas envolvidas na decomposição realizada pelo método EMD.	61
Figura 16 – Monocomponentes produzidos pela decomposição de uma série temporal, utilizando a técnica EMD.	62
Figura 17 – Monocomponentes gerados pela decomposição da série temporal utilizando a técnica EMD.	65
Figura 18 – Exemplo de série temporal produzida pela soma de um componente determinístico e um estocástico.	70
Figura 19 – Comparação entre atratores no espaço-fase.	71
Figura 20 – Estratégia de decomposição do atrator aproximando estados ao esqueleto.	72
Figura 21 – Processo de decomposição pela média dos vizinhos mais próximos.	73
Figura 22 – Exemplo de execução do algoritmo Spring	74
Figura 23 – Comparação entre o atrator do Sistema de Lorenz e o atrator do Mapa Logístico.	76
Figura 24 – Exemplo de execução do algoritmo Spring Time Domain	77
Figura 25 – Exemplo de execução do algoritmo Spring*	80
Figura 26 – Exemplo da diferença entre seno e série contante resultando em soma dos erros igual a zero.	84

Figura 27 – Funcionamento da medida MDDL.	88
Figura 28 – Séries temporais determinísticas e seus respectivos atratores no espaço-fase. Em sequência: Seno, Lorenz, Rössler, Mapa Logístico e Hénon.	90
Figura 29 – Matriz de Correlação calculada para pares de medidas.	91
Figura 30 – Matriz de Avaliação Visual.	95
Figura 31 – Gráfico de Violino com resultados das medidas MAE e MDDL.	98
Figura 32 – Comparação par-a-par do resultado das técnicas.	100
Figura 33 – Séries e atratores dos dados S&P 500.	102
Figura 34 – Séries e atratores dos dados <i>Sunspots</i>	104
Figura 35 – Processo de simulação, construção e mensuração do Circuito de Chua.	105
Figura 36 – Séries e atratores dos dados do Circuito de Chua.	106
Figura 37 – Resultados dos testes de Predição de séries temporais.	108

LISTA DE ALGORITMOS

Algoritmo 1 – Pseudo algoritmo da técnica Spring	75
Algoritmo 2 – Pseudo algoritmo da técnica Spring Time Domain (Spring-TD)	77
Algoritmo 3 – Pseudo algoritmo da técnica Spring*	81

LISTA DE TABELAS

Tabela 1 – Estados do espaço-fase do sistema dinâmico Mapa Logístico.	43
Tabela 2 – Média e desvio padrão da medida MAE.	97
Tabela 3 – Média e desvio padrão da medida MDDL para cada técnica em cada cenário determinístico de teste.	97

LISTA DE ABREVIATURAS E SIGLAS

ARIMA	<i>AutoRegressive Integrated Moving Average</i>
DTW	<i>Dynamic Time Warping</i>
EMD	<i>Empirical Mode Decomposition</i>
EMD-MI	<i>Empirical Mode Decomposition using Mutual Information</i>
EMD-RP	<i>Empirical Mode Decomposition using Recurrence Plot</i>
FNN	<i>False Nearest Neighbors</i>
FT	<i>Fourier Transform</i>
HSA	<i>Hilbert Spectral Analysis</i>
ICMC	Instituto de Ciências Matemáticas e de Computação
IMFs	<i>Intrinsic Mode Functions</i>
MA	<i>Moving Average</i>
MAE	<i>Mean Absolute Error</i>
MDA	<i>Mean Distance from Attractors</i>
MDDL	<i>Mean Distance from the Diagonal Line</i>
MI	<i>Mutual Information</i>
MSE	<i>Mean Squared Error</i>
NYSE	New York Stock Exchange
PCA	<i>Principal Component Analysis</i>
RMSE	<i>Root Mean Squared Error</i>
RP	<i>Recurrence Plot</i>
S&P 500	Standard & Poor's 500
SL	<i>Subspace learning</i>
Spring-TD	Spring Time Domain
SSA	<i>Singular Spectrum Analysis</i>
SVD	<i>Singular Value Decomposition</i>
USP	Universidade de São Paulo
WT	<i>Wavelet Transform</i>

SUMÁRIO

1	INTRODUÇÃO	25
2	ANÁLISE DE SÉRIES TEMPORAIS	33
2.1	Séries temporais	33
2.2	Análise Estatística	37
2.2.1	<i>Modelo Autorregressivo</i>	37
2.2.2	<i>Modelo de médias móveis</i>	37
2.2.3	<i>Modelo autorregressivo e de médias móveis</i>	38
2.2.4	<i>Modelo autorregressivo, integrado e de médias móveis</i>	38
2.2.5	<i>Avaliação do modelo ARIMA</i>	38
2.3	Análise baseada em Sistemas Dinâmicos	39
2.3.1	<i>Conceitos essenciais</i>	39
2.3.2	<i>Reconstrução e desconstrução do espaço-fase</i>	42
2.3.3	<i>Estimação da dimensão de separação</i>	43
2.3.4	<i>Estimação da dimensão embutida</i>	47
2.4	Considerações Finais	49
3	DECOMPOSIÇÃO DE SÉRIES TEMPORAIS	51
3.1	Transformada de Fourier	52
3.2	Transformada Wavelet	54
3.3	Média Móvel	56
3.4	Análise Espectral Singular	57
3.4.1	<i>Estágio de decomposição</i>	58
3.4.2	<i>Estágio de reconstrução</i>	59
3.5	Empirical mode decomposition	59
3.5.1	<i>Empirical mode decomposition using recurrence plot</i>	63
3.5.2	<i>Empirical mode decomposition using mutual information</i>	63
3.6	Métodos de decomposição não linear	66
3.7	Considerações Finais	67
4	ABORDAGENS PROPOSTAS PARA DECOMPOSIÇÃO DE SÉRIES TEMPORAIS	69
4.1	Spring	72
4.2	Spring Time Domain	75

4.3	Spring*	78
4.4	Considerações Finais	79
5	MEDIDAS DE AVALIAÇÃO DE TÉCNICAS DE DECOMPOSIÇÃO	83
5.1	Distância de Minkowski, MAE, MSE e RMSE	83
5.2	Correlação de Pearson	85
5.3	Dynamic Time Warping	86
5.4	Distância Média da Linha Diagonal	86
5.5	Distância média dos Atratores	87
5.6	Bases de Dados Sintéticas	89
5.7	Comparação	91
5.8	Considerações Finais	92
6	COMPARAÇÃO ENTRE TÉCNICAS DE DECOMPOSIÇÃO DE SÉRIES TEMPORAIS	93
6.1	Avaliação Visual	93
6.2	Avaliação Quantitativa	96
6.3	Considerações Finais	99
7	PREDIÇÃO DE SÉRIES TEMPORAIS	101
7.1	Séries Temporais Reais	101
7.1.1	<i>S&P 500</i>	101
7.1.2	<i>Sunspots Mensal</i>	102
7.1.3	<i>Circuito de Chua</i>	103
7.2	Metodologia dos Experimentos e Resultados	106
7.3	Considerações Finais	109
8	CONCLUSÕES	111
8.1	Conclusões Finais	111
8.2	Publicações e Colaborações	112
8.3	Trabalhos Futuros	112
	REFERÊNCIAS	115

INTRODUÇÃO

A evolução tecnológica tem motivado, e em muitos casos permitido, a coleta e a modelagem de grandes volumes de dados produzidos ao longo do tempo. Esses dados, que são tipicamente gerados por processos industriais, operações humanas ou fenômenos da natureza, podem ser organizados na forma de séries temporais (LAU; WENG, 1995; BERNARD; JONES, 1996). Essa organização permite modelagens voltadas para a compreensão desses fenômenos com o intuito de predizê-los, observar suas alterações e controlá-los (ALLIGOOD; SAUER; YORKE, 1996; KANTZ; SCHREIBER, 2004). Considere, por exemplo, a região climática do Sahel, que ocorre sobre a África, entre o deserto do Saara e o norte da Savana Sudanesa. Alterações no Sahel influenciam diretamente no clima de grande parte da África, bem como sobre suas regiões cultiváveis (BRANDT *et al.*, 2014), portanto, ao modelar e compreender o comportamento temporal do Sahel, pode-se tomar ações preventivas sobre a agricultura da região.

Dois principais ramos da ciência buscam modelar e estudar tais fenômenos por meio de séries temporais. O primeiro ramo é Estatístico (MORETTIN; TOLOI, 2006; MONTGOMERY; JENNINGS; KULAHCI, 2015) e considera que séries temporais apresentam influências estocásticas em sua formação. O segundo, baseado em Sistemas Dinâmicos (KANTZ; SCHREIBER, 2004), parte do pressuposto de que toda observação de uma série temporal é produzida somente em função de suas observações anteriores, ou seja, uma regra ou função determinística é responsável por produzir os dados¹.

O ramo Estatístico parte de um modelo pré-definido, tipicamente um dos modelos Auto-Regressivos Integrados de Médias Móveis (do inglês *AutoRegressive Integrated Moving Average* (ARIMA)) propostos por Box e Jenkins (1976), e busca otimizar seus coeficientes a fim de aproximá-los das observações da série temporal em estudo. Em contrapartida, o ramo baseado

¹ Há também a subárea de Sistemas Dinâmicos Estocásticos que utiliza ferramental similar ao ramo aqui denominado de Estatístico.

em Sistemas Dinâmicos utiliza ferramentas para reconstruir a série temporal em um novo espaço, denominado espaço-fase (TAKENS, 1981), o qual desdobra os relacionamentos temporais existentes entre observações (KENNEL; BROWN; ABARBANEL, 1992) para, posteriormente, estimar a regra que produziu tais dados (HEGGER; KANTZ; SCHREIBER, 1999). Observe que o primeiro ramo parte da suposição de um modelo ideal, enquanto o segundo permite encontrar o modelo mais adequado (ALLIGOOD; SAUER; YORKE, 1996).

Experiências prévias justificam que pesquisadores adotem um desses dois ramos para a modelagem de seus dados. Contudo, ao modelar uma série com características predominantemente estocásticas, por meio de ferramentas da área de Sistemas Dinâmicos, tende-se a obter espaços-fase pouco representativos. Logo, as regras geradoras estimadas a partir desses espaços não se adéquam aos dados originais. De maneira complementar, ao modelar uma série predominantemente determinística, por meio de ferramenta Estatístico, tende-se a subestimar as dependências existentes entre observações (RIOS, 2013). O cenário ideal se dá pela adoção de ambas abordagens sobre cada uma dessas influências presentes na série temporal original, ou seja, por meio do emprego do ramo Estatístico sobre as influências estocásticas e do ramo baseado em Sistemas Dinâmicos sobre as determinísticas. Para isso, faz-se necessária a decomposição da série temporal nessas duas influências.

Diversas abordagens foram propostas para decompor série temporais em seus componentes, tais como: Transformada de Fourier (do inglês *Fourier Transform* (FT) – Seção 3.1), Transformada Wavelet (do inglês *Wavelet Transform* (WT) – Seção 3.2), Médias Moveis (do inglês *Moving Average* (MA) – Seção 3.3), Análise Espectral Singular (do inglês *Singular Spectrum Analysis* (SSA) – Seção 3.4), Decomposição de Modo Empírico utilizando Gráficos de Recorrência (do inglês *Empirical Mode Decomposition using Recurrence Plot* (EMD-RP) – Seção 3.5.1) e Decomposição de Modo Empírico utilizando Informação Mútua (do inglês *Empirical Mode Decomposition using Mutual Information* (EMD-MI) – Seção 3.5.2). Em comum, todas estas apresentam um certo grau de deficiência relacionada ao viés adotado para representar o componente determinístico. Deste modo, o atrator² resultante da decomposição é distorcido e, em alguns cenários, tem seu viés aproximado ao comportamento de uma função senoidal mesmo quando a série temporal original não apresenta tal comportamento. Consequentemente, essas abordagens reconstroem atratores não representativos para séries temporais em estudo.

Para exemplificar o problema relativo à abordagem, considere uma série temporal produzida a partir da mistura de um comportamento estocástico (Figura 1a), baseado em uma distribuição de probabilidades Gaussiana (Normal), e um comportamento determinístico (Figura 1b), produzido pelo Mapa Logístico $x_{t+1} = r \times x_t \times (1 - x_t)$, em que x_t determina o estado atual e $r \in \mathbb{R}$ é um parâmetro interno do sistema. Seja a mistura aditiva desses componentes responsável por produzir a série temporal X , ilustrada na Figura 1c.

² Define-se **atrator** como uma nuvem de pontos no espaço-fase que representa o relacionamento entre as observações de uma série temporal em instantes distintos de tempo. Mais informações vide Seção 2.3.1.

Figura 1 – Soma do componente estocástico (a) com o determinístico (b) produz a série temporal (c).

Fonte: Elaborada pelo autor.

Ao aplicar, por exemplo, o algoritmo EMD-RP (Seção 3.5.1) sobre X , dois componentes são obtidos: um estocástico (Figura 2a) e outro determinístico (Figura 2b). Nota-se que o componente determinístico obtido é resultante da soma de funções senoidais. Logo, o desdobramento desse componente no espaço-fase, a fim de obter sua regra geradora, irá produzir um espaço característico de séries temporais senoidais. Esse espaço-fase obtido para séries senoidais apresenta formas elípticas, tais como as ilustradas na Figura 3b.

Contudo, o espaço-fase para o Mapa Logístico forma uma parábola com concavidade para baixo (típico de uma função côncava), tal como ilustrado na Figura 3a. Dessa maneira, a decomposição aplicada pelas atuais abordagens do estado da arte — em particular, neste exemplo, ilustrada pela abordagem EMD-RP — descaracteriza o espaço-fase da série temporal, o que permite concluir que não ocorre a separação ideal entre os componentes estocástico e

(a) Componente estocástico produzido pela decomposição EMD-RP.

(b) Componente determinístico produzido pela decomposição EMD-RP.

Figura 2 – Componentes estocástico (a) e determinístico (b) produzidos pela decomposição da série temporal ilustrada na Figura 1c, utilizando a técnica EMD-RP (Seção 3.5).

Fonte: Elaborada pelo autor.

(a) Formato do atrator esperado.

(b) Atrator produzido para o componente determinístico resultante da técnica EMD-RP.

Figura 3 – O viés senoidal produzido pelas abordagens EMD-RP e EMD-MI distorcem o espaço-fase do componente determinístico. Em (a), tem-se o espaço-fase esperado para o componente determinístico e, em (b), o espaço-fase produzido pela abordagem EMD-RP.

Fonte: Elaborada pelo autor.

determinístico. Essa limitação leva a piores resultados de modelagem e, consequentemente, de predição, uma vez que a regra geradora estimada para o componente determinístico não corresponde à regra de fato utilizada para produzir os dados.

Para contornar tal limitação, as técnicas Lazy e GHKSS (Seção 3.6) se propõem a separar os componentes estocástico e determinístico trabalhando diretamente sobre o atrator da série temporal no espaço-fase. Essas abordagens garantem que a topologia do atrator seja preservada e que o viés original da série sofra o mínimo possível de distorções. Apesar de apresentar bons resultados, tais técnicas apresentam alta sensibilidade a seus hiper-parâmetros e complexidade para configurá-los, ambos deixados a cargo do usuário. Ademais, essas abordagens apresentam limitações para lidar com mudanças na densidade dos estados ao longo do atrator.

Estas limitações motivaram as pesquisas que compõem esta tese de doutorado, proposta com o objetivo de decompor séries temporais em seus respectivos componentes determinísticos e estocásticos preservando, contudo, o viés determinístico original da série. Desse modo, obtém-se modelos mais representativos e, consequentemente, maior acurácia para tarefas de predição.

Neste contexto três técnicas foram desenvolvidas: (i) **Spring** (Seção 4.1), (ii) **Spring Time Domain** (Seção 4.2) e (iii) **Spring*** (Seção 4.3). Em comum, todas operam a etapa de decomposição sobre o espaço-fase a fim de manter a topologia dos atratores e, por conseguinte, o viés do componente determinístico. Diferentemente das técnicas Lazy e GHKSS, a abordagem **Spring**, além da fácil parametrização, adapta-se a diferentes densidades relativas aos estados presentes no espaço-fase, melhorando os resultados da decomposição e das predições. Há, contudo, cenários específicos em que os atratores originais apresentam transições suaves entre seus estados e os resultados produzidos pela **Spring** poderiam, ainda, ser otimizados, dando origem à técnica **Spring Time Domain**. Essa segunda abordagem utiliza splines cúbicas para suavizar transições que foram perturbadas pela presença de componentes estocásticos, levando às melhores separação do componente estocástico e preservação do viés determinístico.

As técnicas **Spring** e **Spring Time Domain** requerem parâmetros definidos pelo usuário para executar o processo de decomposição. Essa tarefa requer um conhecimento profundo sobre o domínio do problema, bem como sobre as características dessas técnicas. Esses fatores motivaram o projeto de uma abordagem sem parametrização, inspirada na EMD-MI, denominada **Spring***, a qual segue a mesma metodologia da **Spring**. Apesar do claro benefício de evitar que usuários realizem parametrizações, a **Spring*** apresenta resultados aquém daqueles produzidos por técnicas tradicionais, superando, no entanto, os resultados obtidos pela EMD-MI. Assim, a **Spring*** apresenta-se como uma opção para auxiliar usuários sem amplos conhecimentos sobre decomposição de séries temporais ou sobre os dados de interesse.

Para comparar as abordagens aqui propostas com as diversas técnicas que compõem o estado da arte, fez-se necessária a definição de critérios a serem utilizados. No entanto, devido à decomposição de séries temporais ser tratada de maneira independente nos mais diversos ramos da ciência, não há uma concordância entre os pesquisadores sobre qual o melhor critério ou forma

de avaliação. Sendo assim, uma análise das medidas e dos testes estatísticos mais usados foi realizada a fim de determinar a melhor metodologia para tais comparações. Os resultados deste estudo guiaram as avaliações no contexto deste doutorado, bem como unificaram as diversas linhas de estudo associadas às melhores metodologias para avaliar técnicas de decomposição.

Neste contexto, diversos testes foram realizados para comparar a **Spring** e suas variações em relação ao estado da arte. Os resultados visuais e quantitativos, a partir de dados sintéticos, mostram que a técnica **Spring** apresenta maior desempenho que as demais abordagens que compõem o estado da arte. Por sua vez, para obter comprovações sobre dados reais, as bases Standard & Poor's 500 (S&P 500), *Sunspots* Mensal e Circuito de Chua foram submetidas aos processos de decomposição e predição. Os resultados confirmaram a hipótese de que ao decompor séries temporais, há uma melhoria significativa no processo de predição quando feito o emprego de **Spring** e suas variantes.

Os capítulos seguintes aprofundam os conceitos abordados na presente seção. O(a) leitor(a) familiarizado(a) com análise e decomposição de séries temporais pode saltar os Capítulos 2 e 3 sem o risco de comprometer o entendimento das principais contribuições dessa tese, cujos capítulos estão organizados da seguinte maneira:

Capítulo 2 – Análise de Séries Temporais – Introduz conceitos relacionados a séries temporais e apresenta os ferramentais utilizados pelas áreas de Estatística e Sistemas Dinâmicos para analisar e compreender o comportamento desse tipo de dado;

Capítulo 3 – Decomposição de Séries Temporais – Apresenta as técnicas do estado da arte, suas principais características e como elas são utilizadas no contexto de decomposição de séries temporais;

Capítulo 4 – Abordagens Propostas para Decomposição de Séries Temporais – Descreve as abordagens propostas neste trabalho para decomposição de séries temporais: **Spring**, **Spring Time Domain** e **Spring***;

Capítulo 5 – Medidas de Avaliação de Técnicas de decomposição – Discute as principais medidas utilizadas para avaliar o resultado de decomposições de séries temporais, comparando-as e determinando quais as melhores medidas para avaliar técnicas no contexto deste trabalho;

Capítulo 6 – Comparação entre Técnicas de Decomposição de Séries Temporais – Detalha e compara os diversos experimentos realizados com as técnicas de decomposição;

Capítulo 7 – Predição de Séries Temporais – Avalia a hipótese deste trabalho, que considera que a decomposição de uma série temporal melhora a predição de novas observações;

Capítulo 8 – Conclusões – Conclui este documento com uma discussão sobre os principais resultados obtidos. Por fim, são elencados os próximos passos e suas relações com o trabalho realizado.

ANÁLISE DE SÉRIES TEMPORAIS

Sendo de relevante aplicação para os mais diversos ramos da ciência, a área de análise de séries temporais fornece ferramentas necessárias para modelar, compreender e prever fenômenos que variam ao longo do tempo, tais como os oriundos da natureza, de processos industriais, de índices da bolsa de valores entre outros (KANTZ; SCHREIBER, 2004; BOX; JENKINS, 1976). Tipicamente, essas análises são realizadas utilizando métodos de duas áreas: Estatística e Sistemas Dinâmicos. Neste capítulo, são abordados ambos ramos e suas principais ferramentas.

2.1 Séries temporais

Define-se uma série temporal X como uma sequência de observações $x_t \in \mathbb{R}^n$ para $n \geq 1$, ordenadas ao longo do tempo (MORETTIN; TOLOI, 2006). Séries temporais podem ser definidas como contínuas, quando suas observações são coletadas em todo instante de tempo em $\mathbb{R}_+$ (Figura 4a), ou discretas, caso em que a aquisição dos dados é feita em intervalos uniformemente espaçados no tempo (Figura 4b). Pode-se, porém, fazer uma amostragem de uma série temporal contínua, a partir de intervalos uniformes de tempo, transformando-a em uma série discreta.

Outra característica intrínseca às séries temporais é relativa ao número de variáveis observadas e coletadas a cada instante de tempo. Se uma única variável é observada, a série é dita unidimensional ou univariada. Quando mais de uma variável é observada, tem-se uma série multidimensional, também denominada multivariada (MORETTIN; TOLOI, 2006). Independente da dimensionalidade, de modo geral, pode-se classificar as séries temporais de acordo com seu (i) determinismo ou estocasticidade, (ii) linearidade e estacionariedade (ISHII; RIOS; MELLO, 2011).

Uma série temporal X , formada por observações $\{x_t, t = 1, \dots, n\}$, é definida como

Figura 4 – Em (a), um exemplo de série temporal contínua que pode ser amostrada em espaços uniformes de tempo, formando uma série temporal discreta como a representada em (b).

Fonte: Elaborada pelo autor.

determinística se uma observação x_t depende única e exclusivamente de uma combinação de observações anteriores. Neste contexto, modelos da área de Sistemas Dinâmicos, especificamente aqueles propostos pela Teoria do Caos (ALLIGOOD; SAUER; YORKE, 1996; KANTZ; SCHREIBER, 2004), são utilizados para o estudo e análise dessas séries, pois produzem melhores resultados. Por outro lado, se uma observação x_t sofre influências de alguma variável aleatória a_t , a série é definida como estocástica. Neste caso, modelos estatísticos, tais como os propostos por Box e Jenkins (1976), tendem a produzir melhores resultados.

Deve-se ressaltar que séries obtidas a partir de fenômenos reais raramente se comportam de maneira puramente determinística. Essa característica se deve à influência de fatores desconhecidos, os quais produzem influências aleatórias tipicamente provenientes da falta de precisão, erro de medição ou falta de conhecimento sobre o fenômeno observado. Assim, em cenários reais, séries temporais são tipicamente formadas por uma mistura desses dois componentes: o determinístico e o estocástico (RIOS, 2013).

As séries temporais com características estocásticas podem ainda ser categorizadas de acordo com sua estacionariedade. Essa organização leva em consideração as relações implícitas entre as observações (ISHII; RIOS; MELLO, 2011). Uma série é dita estacionária quando ela se desenvolve no tempo, mantendo, contudo, média e variância constantes. Essa característica reflete um equilíbrio estável da série. A Figura 5a exemplifica esse cenário e ilustra uma série estacionária com média constante.

Cenários reais, entretanto, costumam apresentar um componente de tendência que caracteriza a não-estacionariedade (MORETTIN; TOLOI, 2006). Por outro lado, há cenários tais como os relacionados a aplicações econômicas e financeiras em que as séries são caracterizadas pelo processo de Passeio Aleatório (do inglês *Random Walk*), ou outros em que séries mensuram temperaturas médias da superfície terrestre (Figura 5b), os quais apresentam tendências lineares, i.e., suas observações flutuam ao redor de uma reta com inclinação positiva ou negativa. Séries oriundas desses processos são conhecidas como não-estacionárias com tendência linear. Outro cenário típico é encontrado no crescimento populacional de colônias de bactérias (Figura 5c), cujas séries apresentam tendências não lineares e suas observações oscilam ao redor de uma média com crescimento exponencial. Estas são conhecidas como séries não-estacionárias com tendência exponencial.

Independentemente das características das séries, as metodologias empregadas para analisá-las buscam modelá-las a fim de estudar perturbações ou alterações comportamentais, compreender os fenômenos geradores ou prever observações. Há dois principais ramos que buscam modelar séries temporais, o primeiro é baseado em ferramentas Estatísticas (SABZIKAR; MCLEOD; MEERSCHAERT, 2019; ROUSSEEUW *et al.*, 2019; JIANG *et al.*, 2018) e o segundo considera conceitos da área de Sistemas Dinâmicos (RIOS, 2013; MFOUMOU; KENMOE; KOFANE, 2019). Tipicamente, a abordagem Estatística aplica ferramentas de autocorrelação sobre uma série temporal com o intuito de encontrar qual dos modelos clássicos da literatura é o mais adequado (tipicamente modelos pertencentes à classe dos modelos autoregressivos integrados de médias móveis, do inglês *Autoregressive Integrated Moving Average*—ARIMA (BOX; JENKINS, 1976)), para, em seguida, estimar seus parâmetros a fim de melhor representar os dados. Em contrapartida, a abordagem baseada em Sistemas Dinâmicos busca reconstruir os dados de uma série temporal em um espaço multidimensional, denominado espaço-fase (TAKENS, 1981), com o objetivo de estimar uma função (também denominada regra geradora) que aproxime o comportamento desses dados. Dessa maneira, a primeira abordagem realiza uma análise inicial para a escolha de um modelo conhecido, enquanto a segunda permite produzir qualquer modelo a partir do espaço-fase.

Ambos os ramos são relevantes e adequados para tipos distintos de séries temporais. O ramo Estatístico produz melhores resultados para séries temporais estocásticas, enquanto o baseado em Sistemas Dinâmicos é melhor para séries temporais determinísticas. Portanto, ambos se complementam e podem ser utilizados em conjunto, tal como proposto por Rios (2013). As seções seguintes apresentam maiores detalhes sobre esses dois ramos empregados na análise de séries temporais.

(a) Série estacionária formada por uma distribuição de probabilidades Normal $N(0, 1)$.

(b) Histórico de temperaturas da superfície terrestre. Série não-estacionária com tendência linear. Dados disponibilizados por [Morice et al. \(2012\)](#).

(c) Crescimento populacional de uma colônia de bactérias sem a limitação de recursos, o que produziria a eliminação de indivíduos ao longo do tempo.

Figura 5 – Exemplos de séries temporais estocásticas em que (a) apresenta uma série estacionária, (b) uma série não-estacionária com tendência linear positiva, e (c) uma série não-estacionária com tendência exponencial. Em vermelho suas respectivas médias, utilizadas para ilustrar a não-estacionariedade.

Fonte: Elaborada pelo autor.

2.2 Análise Estatística

No que se refere à análise de séries temporais por meio do ramo Estatístico, os trabalhos mais relevantes foram desenvolvidos por [Box e Jenkins \(1976\)](#), os quais projetaram uma classe com quatro modelos, denominados modelos “Auto-Regressivos Integrados de Médias Móveis” (do inglês *Autoregressive Integrated Moving Average* – ARIMA). Esses modelos representam as relações de dependência entre as observações de uma série temporal sob a influência de processos estocásticos, e permitem representar séries temporais estacionárias e não-estacionárias.

A abordagem de [Box e Jenkins \(1976\)](#) modela uma série temporal em três etapas. Primeiramente, identifica-se o modelo da classe ARIMA mais adequado para representar as observações da série temporal. Em seguida, realiza-se a fase de adaptação dos parâmetros desse modelo em relação às observações da série. Por fim, verifica-se a qualidade do modelo e, caso não seja adequado, todo o processo é novamente realizado. É essencial a obtenção de modelos representativos com o menor número de parâmetros possível, respeitando o princípio da parcimônia ([SOKLAKOV, 2002](#)), evitando assim o superajustamento (do inglês, *overfitting*) do modelo às observações disponíveis. A classe ARIMA é composta pelos modelos apresentados nas seções seguintes.

2.2.1 Modelo Autorregressivo

O modelo autorregressivo de ordem p , denominado $AR(p)$, é indicado para séries temporais estacionárias. De acordo com esse modelo, o valor da observação atual x_t é definido a partir de um sistema linear finito de observações prévias somado a um ruído aleatório a_t , conforme definido na Equação (2.1), em que os valores $\phi_1, \phi_2, \dots, \phi_p$ correspondem aos coeficientes do modelo ou pesos para as observações prévias, e o termo p corresponde à ordem do modelo, ou seja, ao número de observações passadas que devem ser levadas em consideração.

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + a_t \quad (2.1)$$

2.2.2 Modelo de médias móveis

O modelo de médias móveis de ordem q , denominado $MA(q)$, pressupõem que a observação atual é formado pela média ponderada das q observações anteriores. Assim, o valor da observação atual x_t é definido por um sistema linear finito das observações prévias x_t , conforme definido na Equação (2.2). $\theta_1, \theta_2, \dots, \theta_q$ correspondem aos coeficientes do modelo ou pesos para os valores prévios, cuja soma é igual a 1, e o termo q determina o número de observações passadas que devem ser levadas em consideração.

$$x_t = x_t + \theta_1 x_{t-1} + \theta_2 x_{t-2} + \dots + \theta_q x_{t-q} \quad (2.2)$$

2.2.3 Modelo autorregressivo e de médias móveis

O modelo autorregressivo de médias móveis de ordem p, q , denominado $ARMA(p, q)$, é formado pela união do modelo Autoregressivo $AR(p)$ e o de média móveis $MA(q)$. É utilizado para séries cujo valor de alguma variável no instante t é definido em função de valores defasados desta mesma variável em instantes passados (MORETTIN; TOLOI, 2006). De acordo com este modelo, o valor da observação atual x_t é definido a partir de um sistema linear finito de observações prévias – $AR(p)$ – somado a um outro sistema linear finito dos valores prévios de ruídos aleatórios a_t – $MA(q)$ –, conforme definido na Equação (2.3), em que ϕ e θ são os coeficientes do modelo autorregressivo e de médias móveis, respectivamente.

$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.3)$$

2.2.4 Modelo autorregressivo, integrado e de médias móveis

O modelo autorregressivo, integrado e de médias móveis $ARIMA(p, d, q)$ é utilizado para representar séries temporais não estacionárias com tendência linear. Define-se uma série estacionária w_t como sendo a diferença de ordem d da série não estacionária x_t , tal como definido na Equação (2.4).

$$w_t = \Delta^d x_t \quad (2.4)$$

Dessa forma, x_t é definida como uma integral de w_t e, portanto, diz-se que x_t segue um modelo autorregressivo, *integrado*, de médias móveis de acordo com a Equação 2.5, em que ϕ e θ são os coeficientes do modelo autorregressivo e de médias móveis, respectivamente.

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q} \quad (2.5)$$

2.2.5 Avaliação do modelo ARIMA

Para encontrar os parâmetros do modelo ARIMA que melhor se adaptam à série temporal em estudo, pode-se utilizar os estimadores de mínimos quadrados ou de máxima verossimilhança (BOWERMAN, 1987). Independentemente da abordagem adotada, os parâmetros estimados produzem um modelo que precisa ser avaliado com o intuito de verificar se esse representa, satisfatoriamente, os dados em estudo. Para isso, a metodologia utilizada consiste em analisar os resíduos. Para que o modelo esteja correto, os resíduos devem ser independentes e identicamente distribuídos (i.i.d.), seguindo uma distribuição de probabilidades Normal com média zero e variância σ^2 , i.e., $\mathcal{N}(0, \sigma^2)$.

Neste contexto, Box e Pierce (1970) propuseram utilizar um teste de hipótese para verificar as autocorrelações dos resíduos estimados. Mais tarde, Ljung e Box (1978) propuseram uma pequena modificação sobre a estatística do teste, tornando-o mais genérico e produzindo

melhores resultados. As hipóteses para o teste são:

$$\begin{cases} H_0 : \text{"Os resíduos são i.i.d."} \\ H_1 : \text{"Os resíduos não são i.i.d."} \end{cases} \quad (2.6)$$

em que H_0 indica que as correlações da população, da qual a amostra é colhida, são iguais a zero, i.e., quaisquer correlações observadas nos dados resultam da aleatoriedade do processo de amostragem. Se o modelo for apropriado, a estatística do teste definida pela Equação (2.7), em que K é o número de defasagens tomadas e $\hat{r}_j$ é a autocorrelação estimada da série, terá aproximadamente uma distribuição de probabilidades Chi-quadrado χ^2 com $K - p - q$ graus de liberdade, em que p e q são as ordens do modelo ajustado. Assim, rejeita-se H_0 se $Q > \chi^2_{1-\alpha, K-p-q}$ com um nível de significância α .

$$Q(k) = n(n-2) \sum_{j=1}^K \frac{\hat{r}_j^2}{(n-j)} \quad (2.7)$$

2.3 Análise baseada em Sistemas Dinâmicos

A análise de séries temporais utilizando ferramental do ramo de Sistemas Dinâmicos é importante para compreender e prever fenômenos com sazonalidade (recorrências) e tendências (ALLIGOOD; SAUER; YORKE, 1996). De modo geral esse ramo se preocupa em encontrar relações entre as observações temporais a fim de determinar a função ou regra geradora da série temporal. Para isso, foram desenvolvidos diversos algoritmos que tomam como base o método dos mínimos quadrados e utilizam funções de base radial ou polinomiais (HEGGER; KANTZ; SCHREIBER, 1999). Esses algoritmos trabalham sobre um espaço multidimensional denominado espaço-fase que é obtido por meio do emprego do Teorema de Imersão proposto por Takens (1981).

Esse espaço permite desdobrar as dependências temporais entre as observações, mapeando assim suas relações na forma de estados. Contudo, antes de empregar tal teorema, deve-se estimar a dimensão de separação τ , que determina o atraso entre observações, e a dimensão embutida m , que corresponde ao número de observações utilizadas para compor um estado.

2.3.1 Conceitos essenciais

Define-se um sistema dinâmico como um conjunto de possíveis estados em que uma regra determina o estado presente com relação aos estados passados. Esta regra, por sua vez, necessita ser determinística, o que implica em determinar o estado presente unicamente baseado nos estados passados (ALLIGOOD; SAUER; YORKE, 1996). Assim, cada estado, definido por uma variável $s_t \in \mathbb{R}^m$, contém toda a informação necessária para que uma regra ou função geradora f transite, sem dúvida, do sistema do estado s_t para um novo s_{t+1} , ou seja $s_{t+1} = f(s_t)$. O conjunto desses estados forma o espaço $\mathbb{S}^n \subseteq \mathbb{R}^m$, i.e., $s_t \in \mathbb{S}^n, t = 0, \dots, t, \dots, n$. Tais estados evoluem no sistema dinâmico de acordo com a regra f .

Dessa maneira, dado o estado inicial s_0 , a primeira iteração $f(s_0)$ resultará no estado s_1 , a segunda $f(s_1) = f(f(s_0)) = f^2(s_0) = s_2$ e a t -ésima iteração $f(s_{t-1}) = f^t(s_0) = s_t$. Por exemplo, considere $s_0 = 2$ e $f(\hat{x}) = \hat{x}^2$, a primeira iteração pode ser calculada como $f(2) = 4$, a segunda iteração como $f(f(2)) = f(4) = 16$, a terceira iteração $f^3(2) = f(16) = 256$ e assim sucessivamente. Ao fim, após t iterações, obtém-se o conjunto de estados $\{s_0, f(s_0), f^2(s_0), \dots, f^t(s_0)\}$. Este conjunto, denominado órbita ou trajetória, pode ser visto como uma série temporal se organizado ao longo do tempo.

Existem situações particulares em que $f(s_t) = s_{t+1} = s_t$, i.e., o estado atual produzido pela função geradora é o mesmo que o estado passado. Tal estado, denotado por $\hat{p}_1$, é dito como 1-periódico e tem $f(s_t) = s_t$. Há ainda cenários em que a função geradora aplicada sobre o estado s_t leva ao estado s_{t+1} e, quando aplicada sobre s_{t+1} , leva de volta ao estado s_t . Nesta situação, em que $f(s_t) = s_{t+1}$ e $f(s_{t+1}) = s_t$, tem-se um ponto fixo com grau 2 de periodicidade (2-periódico) denotado por $f^2(s_t) = s_t$. É possível generalizar tal periodicidade tal que um ponto k -periódico é definido por $f^k(s_t) = s_t$. A análise desses comportamentos é importante para mapear recorrências e padrões no fenômeno em estudo.

Além disso, o estudo de pontos fixos permite encontrar regiões do espaço-fase para onde órbitas convergem, as quais são denominadas regiões de atração ou, mais comumente, atratoras. Considere um estado m -dimensional $s_t = (x_1, x_2, \dots, x_m) \in \mathbb{S}^n$ e uma ε -vizinhança de s_t denominada $N_\varepsilon(s_t)$. Essa vizinhança determina o subconjunto de estados com distância menor que ε em relação ao estado s_t , ou seja, $N_\varepsilon(s_t) = \{\hat{v} \in \mathbb{S}^n \mid |\hat{v} - s_t| < \varepsilon\}$. Se $\lim_{k \rightarrow \infty} f^k(\hat{v}) = s_t$, i.e., se todos os pontos $\hat{v}$ pertencentes à vizinhança $N_\varepsilon(s_t)$ convergem para o estado s_t ao longo do tempo, s_t é considerado atrator e a vizinhança $N_\varepsilon(s_t)$ é denominada bacia de atração (ALLIGOOD; SAUER; YORKE, 1996).

De modo geral, o comportamento de um sistema dinâmico pode ser afetado por dois motivos: os valores atribuídos aos parâmetros da função geradora e/ou o valor do estado inicial. Considere, por exemplo, o sistema dinâmico unidimensional denominado Mapa Logístico, cuja função geradora é definida por $f(x_t) = r \times x_t \times (1 - x_t)$, em que x_t determina o estado atual e $r \in \mathbb{R}$ é um parâmetro interno do sistema. Ao fixar o estado inicial em $x_t = 0,5$ e produzir duas séries com $r = 2,8$ e $r = 3,8$, após 300 iterações, é possível notar uma diferença significativa entre o comportamento das séries (Figura 6a e Figura 6b). Para o cenário com $r = 2,8$ é possível observar que após 22 iterações (linha vermelha na Figura 6a) a série converge para o ponto fixo 0,64. Este mesmo comportamento não acontece no cenário com $r = 3,8$ (Figura 6b), o qual pode ser confundido com o comportamento de um sistema estocástico.

Considere, ainda, o mesmo sistema dinâmico (Mapa Logístico), fixando, porém, o valor do parâmetro do sistema em $r = 3,8$. Ao variar o estado inicial de tal maneira que a distância entre eles seja infinitesimal, para exemplificar seja $x_t = 0,5$ e $x'_t = 0,50000001$, a diferença entre as órbitas geradas cresce consideravelmente com a evolução do tempo (Figura 6c). A essa sensibilidade às condições iniciais dá-se o nome de caoticidade (OTT; SAUER; YORKE,

(a) Mapa Logístico com $r = 2,8$ e $x_0 = 0,5$. (b) Mapa Logístico com $r = 3,8$ e $x_0 = 0,5$. (c) Duas séries temporais formadas pelo Mapa Logístico com $r = 3,8$. Em preto a série com estado inicial $x_0 = 0,5$ e em vermelho a série formada com estado inicial $x_0 = 0,50000001$. É possível notar como as séries são sensíveis ao estado inicial pois divergem consideravelmente ao longo das iterações.

Figura 6 – Sistema dinâmico unidimensional denominado Mapa Logístico com função geradora $f(x_t) = r \times x_t \times (1 - x_t)$. Em (a), o parâmetro $r = 2,8$ produz um comportamento convergente para um ponto fixo (linha vermelha). Em (b), o parâmetro $r = 3,8$ produz efeitos bruscos de transição (típicos de altas frequências), os quais podem parecer provenientes de comportamento estocástico. Em (c), fixa-se o valor de $r = 3,8$ mas se varia o valor do estado inicial de tal maneira que a distância entre eles seja infinitesimal, obtendo-se duas séries que divergem significativamente ao decorrer do tempo (essa sensibilidade à condição inicial é característica de uma série caótica).

Fonte: Adaptada de [Guzzo \(2014, Página 23\)](#).

1994). Todavia, é possível reconstruir essa mesma série em um espaço-fase cujo desdobramento representa as relações entre observações ao longo do tempo. Conforme ilustrado na Figura 7, esse novo espaço mantém o formato do atrator independentemente do estado inicial.

Figura 7 – espaço-fase formado pelas observações da série temporal formada pelo Mapa Logístico. Em preto são apresentados os estados que formam o atrator da série com estado inicial $x_0 = 0,5$ e, em vermelho, o atrator da série formada com estado inicial $x'_0 = 0,50000001$. É possível notar que, independentemente do comportamento caótico da série, o formato do atrator no espaço-fase é preservado.

Fonte: Elaborada pelo autor.

2.3.2 Reconstrução e desconstrução do espaço-fase

Para a reconstrução do espaço-fase, emprega-se o Teorema de Imersão proposto por Takens (1981) sobre a série temporal $X = \{x_0, \dots, x_t, \dots, x_n\}$. Este desdobra as observações x_t da série, produzindo o espaço de estados m -dimensional $\mathbb{S}^n \subseteq \mathbb{R}^m$. A dimensão m , denominada dimensão embutida, determina qual é o número de amostras de X necessárias para compor um estado, ou seja, o estado s_t é composto por m observações de X de modo que a formulação para o Teorema de Imersão seja respeitada, i.e., $s_t = (x_t, x_{t+\tau}, \dots, x_{t+(m-1)\tau})$. Por sua vez, o atraso τ entre observações é denominado dimensão de separação. Em suma, o valor de τ determina o tamanho do “salto”, ou atraso temporal, a ser considerado entre observações da série temporal para se formar um estado s_t do espaço-fase.

Considere o exemplo do sistema dinâmico produzido pelo Mapa de Logístico com $r = 3,8$ e $x_0 = 0,5$. Ao utilizar a dimensão de separação $m = 4$ e dimensão embutida $\tau = 3$, são produzidos os estados presentes na Tabela 1. É possível notar o deslocamento de tamanho τ entre as séries que formam as dimensões do espaço. Devido à essa diferença, o espaço-fase tem exatamente $n - \tau(m - 1)$ elementos, em que t é o tamanho da série temporal.

Devido à característica do espaço-fase em que cada uma das dimensões contém uma subsérie deslocada no espaço tempo, o processo de retorno, desconstrução ou transformação do

espaço-fase para o espaço tempo é trivial. Para isso, qualquer uma das colunas da Tabela 1 pode ser selecionada e os seus valores tratados como amostras da série.

Tabela 1 – Estados do espaço-fase do sistema dinâmico Mapa Logístico, com $r = 3,8$, produzido pelo Teorema de Imersão proposto por [Takens \(1981\)](#) com dimensão de separação $\tau = 3$ e dimensão embutida $m = 4$. Em destaque na tabela, os quadros mostram o efeito da dimensão de separação τ entre as dimensões do atrator, evidenciando o deslocamento dos dados entre as dimensões.

Estados do espaço-fase	Dimensão embutida $m = 4$			
	$\{x_0, \dots, x_{n-3\tau}\}$	$\{x_\tau, \dots, x_{n-2\tau}\}$	$\{x_{2\tau}, \dots, x_{n-\tau}\}$	$\{x_{3\tau}, \dots, x_n\}$
s_1	0,500000	0,562095	0,672557	0,948654
s_2	0,950000	0,935348	0,836851	0,185096
s_3	0,180500	0,229794	0,518819	0,573174
s_4	0,562095	0,672557	0,948654	0,929653
s_5	0,935348	0,836851	0,185096	0,248514
s_6	0,229794	0,518819	0,573174	0,709668
s_7	0,672557	0,948654	0,929653	0,782949
s_8	0,836851	0,185096	0,248514	0,645771
s_9	0,518819	0,573174	0,709668	0,869254
s_{10}	0,948654	0,929653	0,782949	0,431877
s_{11}	0,185096	0,248514	0,645771	0,932365
s_{12}	0,573174	0,709668	0,869254	0,239630
s_{13}	0,929653	0,782949	0,431877	0,692388
s_{14}	0,248514	0,645771	0,932365	0,809350
s_{15}	0,709668	0,869254	0,239630	0,586351
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

As dimensões embutida e de separação não podem ser escolhidas arbitrariamente ([KENNEL; BROWN; ABARBANEL, 1992](#); [FRASER; SWINNEY, 1986](#)). Seus valores devem ser bem estimados, pois a superestimação ou a subestimação produzem estados pouco significativos para o espaço-fase. As seções seguintes apresentam os principais métodos de estimação para essas dimensões.

2.3.3 Estimação da dimensão de separação

A dimensão de separação τ determina o atraso entre as observações da série temporal que compõem um estado do espaço-fase. A escolha arbitrária dessa dimensão pode gerar estados pouco significativos, diminuindo assim a qualidade do espaço-fase construído. Esse problema é agravado em séries ruidosas, como é o caso de séries temporais provenientes de fenômenos reais ([KANTZ; SCHREIBER, 2004](#)). Assim, é importante se dispor de um método que consiga estimar o valor mais adequado para a dimensão de separação ([OTT; SAUER; YORKE, 1994](#)).

[Fraser e Swinney \(1986\)](#) propuseram o emprego da técnica de Informação Mútua (do inglês *Mutual Information* (MI)) para estimar a dimensão de separação. Essa é capaz de mensurar quanto de informação uma variável aleatória tem (ou representa) sobre qualquer outra. A MI é

igual a zero para cenários em que as variáveis são estatisticamente independentes. A metodologia consiste em calcular MI sobre uma sequência de observações, defasadas em τ instantes de tempo, escolhendo como dimensão de separação o primeiro mínimo local encontrado. Essa escolha se dá por corresponder a um atraso temporal suficiente para o desdobramento de estados significativos no espaço-fase. Esse processo visa permitir que cada estado, ou ponto no espaço-fase, contenha toda a informação necessária e, portanto, não apresente ambiguidades para evoluir no tempo, a partir de uma regra geradora.

A fim de ilustrar o cenário descrito, considere o atrator do Sistema de Lorenz (Figura 8a) formado pelo sistema de equações diferenciais:

$$\begin{cases} \frac{\partial x}{\partial t} = \sigma(y - x) \\ \frac{\partial y}{\partial t} = x(\rho - z) - y \\ \frac{\partial z}{\partial t} = xy - \beta z \end{cases}$$

em que a tríade (x, y, z) define a posição espacial de cada estado em $\mathbb{R}^3$; e $\frac{\partial x}{\partial t}$, $\frac{\partial y}{\partial t}$ e $\frac{\partial z}{\partial t}$ são relacionados ao deslocamento temporal dos respectivos estados no instante de tempo t ; e σ , β e ρ são parâmetros do sistema.

Utilizando somente o componente $\frac{\partial x}{\partial t}$, obtém-se uma série temporal (Figura 8b) que pode ser reconstruída no espaço-fase, por meio do Teorema proposto por [Takens \(1981\)](#), conforme ilustrado na Figura 8a. Ao se aplicar a técnica de Informação Mútua sobre essa série, variando contudo o atraso de tempo τ , obtém-se o resultado ilustrado na Figura 8c. Neste caso, o primeiro mínimo local é encontrado para um atraso de tempo $\tau = 3$ (indicado em vermelho na figura), de acordo com [Fraser e Swinney \(1986\)](#). É possível observar a existência de um segundo mínimo local em $\tau = 8$ (assim como há diversos outros), cujo valor de informação mútua é significativamente menor, porém esses e outros são desconsiderados por essa abordagem.

Cabe aqui ressaltar que a qualidade do atrator no espaço-fase está diretamente relacionada à dimensão de separação τ estimada. Caso o valor encontrado seja muito pequeno, as observações x_k e $x_{k-\tau}$ são muito similares, produzindo estados pouco significativos. Ademais, caso o valor estimado seja muito grande, a correlação entre os estados pode ser nula, gerando também estados pouco significativos. Essa situação pode ser observada na Figura 9. Ao se utilizar $\tau = 0$ para efetuar a imersão da série no espaço-fase (Figura 9a), cada estado contém apenas informações relativas à própria observação, formando um atrator linha. Este atrator não contém informações pertinentes ao estudo do comportamento da série, pois não representa relações entre observações.

De maneira similar, ao utilizar um valor demasiadamente grande para a dimensão de separação, tal como $\tau = 182$ (Figura 9d), obtém-se baixa correlação entre os instantes de tempo, o que distorce por completo o atrator no espaço-fase. Essa distorção provoca emaranhados nas trajetórias dos estados, dificultando a análise e determinação de próximos estados para o sistema dinâmico. Neste caso em particular, os melhores cenários ocorrem para $\tau = 3$ e $\tau = 8$. A abordagem proposta por [Fraser e Swinney \(1986\)](#) sugere o primeiro mínimo local, i.e., $\tau = 3$.

(a) Atrator do Sistema de Lorenz em sua forma bidimensional (considerando apenas as séries resultantes de $\frac{\partial x}{\partial t}$ e $\frac{\partial z}{\partial t}$). O estado inicial do sistema foi determinado na posição $(x, y, z) = (-14, -13, 47)$ com parâmetros $\sigma = 10$, $\beta = 8/3$ e $\rho = 28$.

(b) Série temporal formada pelos valores de $\frac{\partial x}{\partial t}$ para o atrator do Sistema de Lorenz.

(c) Resultado da técnica de Informação Mútua aplicada ao atrator do Sistema de Lorenz para estimar a dimensão de separação. Em azul o primeiro mínimo local que, de acordo com [Fraser e Swinney \(1986\)](#), é o valor da dimensão de separação. Em vermelho, o segundo mínimo estimado em $\tau = 8$ e o mínimo em $\tau = 182$, os quais são desconsiderados por essa abordagem.

Figura 8 – Estimação da dimensão de separação para o atrator do Sistema de Lorenz. Em (a), o atrator do Sistema de Lorenz em sua forma bidimensional. Em (b), a série temporal formada pela equação diferencial $\frac{\partial x}{\partial t}$. Em (c), o resultado do cálculo de Informação Mútua para diversos atrasos de tempo τ .

Fonte: Elaborada pelo autor.

Figura 9 – Atratores do Sistema de Lorenz imersos no espaço-fase, variando-se a dimensão de separação τ . Em (a), $\tau = 0$, em (b), $\tau = 3$, em (c), $\tau = 8$ e, em (d), $\tau = 182$. É possível notar que, para $\tau = 0$, os estados correspondem às observações e, por esse motivo, eles não são significativos. Em contrapartida, para $\tau = 182$, os estados também são pouco significativos, pois a correlação entre eles é nula. O ideal se dá pela adoção de um valor intermediário $\tau = 3$ ou $\tau = 8$.

Fonte: Elaborada pelo autor.

2.3.4 Estimação da dimensão embutida

O número de observações necessárias para compor um estado deve ser corretamente estimado a fim de garantir o desdobramento do espaço-fase. Caso o valor utilizado seja menor que o necessário, o espaço-fase resultante será caracterizado por trajetórias emaranhadas, i.e., irá ocorrer um grande número de cruzamentos das trajetórias entre os estados. Conforme o valor da dimensão embutida m aumenta, o espaço-fase se desdobra e os cruzamentos tendem a desaparecer até que se atinja o valor ideal m_E . Isso significa que se o valor da dimensão embutida, utilizado para o cálculo do espaço-fase, for menor que o ideal ($m < m_E$), o espaço gerado irá conter muitas trajetórias que se cruzam. Caso seja utilizado um valor maior que o ideal ($m > m_E$), o espaço gerado será matematicamente equivalente ao espaço produzido pela utilização do valor ideal m_E , pois o atrator já foi totalmente desdobrado (KENNEL; BROWN; ABARBANEL, 1992).

No entanto, do ponto de vista computacional, quanto maior o valor utilizado para m , maior a complexidade para analisá-lo, o que resulta em aumento considerável de tempo de processamento, desencorajando o uso de grandes dimensões embutidas arbitrárias. Assim, pelo princípio da parcimônia, é interessante obter o menor valor suficiente para a dimensão embutida.

Para exemplificar o cenário anteriormente descrito, considere o sistema dinâmico formado pelo Sistema de Rössler, definido pelas equações diferenciais:

$$\begin{cases} \frac{\partial x}{\partial t} = -y - z \\ \frac{\partial y}{\partial t} = x + ay \\ \frac{\partial z}{\partial t} = b + z(x - c) \end{cases},$$

em que a tríade x, y e z define a posição espacial de cada estado em $\mathbb{R}^3$, $\frac{\partial x}{\partial t}$, $\frac{\partial y}{\partial t}$ e $\frac{\partial z}{\partial t}$ correspondem ao deslocamento temporal dos respectivos estados no instante de tempo t ; e a, b e c são parâmetros do sistema. Dado o estado inicial $\hat{x}_0 = (0, 0, 0)$, os parâmetros internos do sistema $a = 0,2$, $b = 0,2$ e $c = 5,7$ e a dimensão de separação $\tau = 51$ (calculada utilizando Informação Mútua – vide Seção 2.3.3). Ao utilizar a dimensão embutida $m = 2$, obtém-se o atrator ilustrado na Figura 10a, o qual contém muitos cruzamentos de trajetórias, indicando que a dimensão embutida não é suficiente para desdobrar as relações entre os estados. Ao se utilizar a dimensão embutida $m = 3$ (Figura 10b), as trajetórias antes emaranhadas deixam de se cruzar, devido à nova dimensão ser mais adequada e, neste caso, suficiente (KANTZ; SCHREIBER, 2004).

Para estimar a melhor dimensão embutida, Kennel, Brown e Abarbanel (1992) propuseram o método de Falsos Vizinhos Mais Próximos (do inglês *False Nearest Neighbors* (FNN)). Esse método varia o valor da dimensão embutida a cada iteração com o intuito de calcular o vizinho mais próximo de cada estado em um espaço-fase candidato (espaço em análise). Caso o vizinho mais próximo encontrado na iteração i não seja mantido na iteração seguinte, ele é então considerado um falso vizinho. O valor estimado da dimensão embutida é encontrado quando a porcentagem de falsos vizinhos for menor que um limiar (tipicamente 10% (KENNEL;

(a) Atrator do Sistema de Rössler desdobrado com $m = 2$. (b) Atrator do Sistema de Rössler desdobrado com $m = 3$.

Figura 10 – Atratores do Sistema de Rössler. Em (a), o atrator foi imerso em um espaço-fase bidimensional. É possível perceber que existe uma grande quantidade de cruzamentos nas trajetórias do atrator. Em (b), os cruzamentos são desdobrados na terceira dimensão, melhorando significativamente a compreensão das trajetórias e, conseqüentemente, do fenômeno ao qual esses dados estão associados.

Fonte: Elaborada pelo autor.

BROWN; ABARBANEL, 1992)), o qual sugere que o valor da dimensão embutida é suficiente para desdobrar o espaço-fase. Ao aplicar FNN sobre a série resultante do Sistema de Rössler (apresentada no exemplo anterior – Figura 10), obtém-se o resultado apresentado na Figura 11. A partir da análise dessa figura, pode-se concluir que o primeiro valor de dimensão embutida menor que o limiar padrão de 10% é dado por $m = 3$, o qual corrobora as imersões ilustradas nas Figuras 10a e 10b.

Figura 11 – Gráfico de percentual de falsos vizinhos mais próximos em termos da dimensão embutida, a qual varia de 1 a 10. O primeiro valor para a dimensão embutida que apresentar porcentagem menor que certo limiar (neste caso 10% – caracterizado pela linha vermelha) é escolhido como o parâmetro ideal de acordo com a abordagem proposta por Kennel, Brown e Abarbanel (1992).

Fonte: Elaborada pelo autor.

2.4 Considerações Finais

A análise de séries temporais é imprescindível para compreender e prever fenômenos que variam ao longo do tempo. Nesse contexto, este capítulo apresentou a definição formal de séries temporais e as principais técnicas de análise, provenientes dos ramos Estatístico e de Sistemas Dinâmicos. Esses conceitos formam o arcabouço teórico necessário para desenvolvimento e compreensão deste trabalho de doutorado. Em particular, serão consideradas neste estudo séries temporais discretas univariadas que serão decompostas em seus respectivos componentes estocástico e determinístico. Essa decomposição se faz necessária para que cada componente seja tratado com as metodologias de suas respectivas áreas descritas neste capítulo. Espera-se assim melhorar a qualidade dos modelos gerados e, conseqüentemente, a compreensão sobre os fenômenos em estudo.

DECOMPOSIÇÃO DE SÉRIES TEMPORAIS

Conhecimentos prévios fazem com que pesquisadores adotem abordagens Estatísticas ou baseadas em Sistemas Dinâmicos para analisar séries temporais (RIOS, 2013). Essa opção única tende a causar problemas, pois tipicamente subestima influências estocásticas ou determinísticas presentes nos dados. Por exemplo, ao modelar uma série predominantemente estocástica utilizando somente o ferramental da área de Sistemas Dinâmicos, tende-se a obter espaços-fase pouco representativos, cujas recorrências não podem ser analisadas. De maneira complementar, ao modelar uma série predominantemente determinística somente com ferramental Estatístico tende-se a desconsiderar recorrências típicas do componente determinístico. Nesse contexto, o cenário ideal se dá pela adoção de ambas abordagens sobre cada uma de suas influências (ou componentes) específicas presentes na série temporal (DUARTE *et al.*, 2019). No entanto, para permitir esta análise, faz-se necessária a decomposição da série temporal.

Há na literatura diversas abordagens para a decomposição de séries temporais tais como Transformada de Fourier (FT) (BRACEWELL, 1986), Transformada Wavelet (WT) (GRAPS, 1995) e Análise Espectral Singular (SSA) (GOLYANDINA, 2001). Essas três decomposições operam de forma linear sobre séries temporais, prejudicando as características (aqui resumidamente denominadas vieses ou influências) não lineares presentes. Cabe ressaltar que muitas das séries que têm despertado o interesse por parte da comunidade científica apresentam componentes não lineares (KANTZ; SCHREIBER, 2004). Além disso, essas decomposições não consideram recorrências entre observações, o que prejudica na reconstrução de atratores e obtenção de regras (ou funções) geradoras desses dados.

Essas limitações motivaram Huang *et al.* (1998) a projetar o método *Empirical Mode Decomposition* (EMD) para decompor séries temporais lineares e não-lineares em um número finito de monocomponentes representativos. Essa proposta motivou Rios e Mello (2013) a desenvolver duas novas abordagens de decomposição a fim de separar influências determinísticas e estocásticas de séries temporais, as quais são denominadas *Empirical Mode Decomposition*

using Recurrence Plot (EMD-RP) e Empirical Mode Decomposition using Mutual Information (EMD-MI). Apesar do sucesso das abordagens de Rios e Mello (2013), o fato da EMD operar somente sobre o domínio temporal faz com que ela prejudique o viés determinístico dos dados, que será destacado posteriormente.

Assim, utilizando o teorema proposto por Takens (1981), as técnicas Lazy (SCHREIBER, 1993) e GHKSS (GRASSBERGER *et al.*, 1993)¹ imergem as séries temporais no espaço-fase onde trabalham a decomposição a fim de manter o formato topológico do atrator. Esta estratégia, além de produzir bons resultados, preserva o viés determinístico dos dados. Entretanto, o número fixo de vizinhos adotados e a dificuldade de encontrar e entender os parâmetros diminuem sua performance e dificulta sua utilização por pesquisadores de outras áreas. Neste capítulo, são introduzidas essas abordagens para a decomposição de séries temporais, pois constituem base comparativa ao trabalho proposto neste doutorado.

3.1 Transformada de Fourier

A Transformada de Fourier (do inglês *Fourier Transform* – FT) é uma técnica utilizada para decompor um sinal, ou uma série temporal, por uma soma de diversas funções senoidais (BRACEWELL, 1986). Essa habilidade tem sido explorada em vários domínios de aplicação, incluindo a separação de componentes determinísticos e estocásticos quando ocorrem em diferentes intervalos de frequência. Em aplicações do mundo real, pesquisadores empregam a versão discreta da FT, produzindo um conjunto de coeficientes complexos baseados no sinal de entrada. Esses coeficientes correspondem a funções sinusoidais, seguindo a fórmula de Euler (Equação (3.1)) em que j é a frequência associada ao componente sinusoidal.

$$f_j = \frac{1}{n} \sum_{k=0}^{n-1} x_k e^{-\frac{2\pi i}{n} jk} \quad j = 0, \dots, n-1 \quad (3.1)$$

A metodologia de separação de séries temporais em componentes estocásticos e determinísticos assume que coeficientes complexos com maiores magnitudes estão associados ao componente determinístico, uma vez que devem refletir os elementos de dados mais relevantes. Consequentemente, os coeficientes com magnitudes mais baixas são associados ao comportamento estocástico (que em muitos cenários é tido como ruído). Estes coeficientes são zerados e removidos do sinal, restando apenas o componente determinístico que é reconstruído no domínio do tempo (CANDAN; KUTAY; OZAKTAS, 2000). Caso o componente estocástico seja necessário, basta subtrair o componente determinístico da série temporal original.

A fim de ilustrar este processo de decomposição, considere a série temporal mostrada na Figura 12a, que é a soma de uma série determinística senoidal com uma série estocástica que segue uma distribuição Normal $\mathcal{N}(0; 0,01^2)$. Ao calcular a magnitude dos coeficientes comple-

¹ Ambos disponíveis no software Tisean (HEGGER; KANTZ; SCHREIBER, 1999).

(a) Série temporal formada pela adição de um componente determinístico senoidal e um componente estocástico com função de distribuição de probabilidades Normal $\mathcal{N}(0, 0.01^2)$.

(b) Magnitudes dos coeficientes resultantes da FT.

(c) Todos os coeficientes pouco representativos são zerados, filtrando assim o componente estocástico.

(d) Os coeficientes são imersos de volta no espaço temporal e, como pode ser comparado com (a), o componente estocástico foi de fato removido da série temporal.

Figura 12 – Processo de decomposição utilizando a FT. Em (a), tem-se uma série temporal formada pela adição de uma série senoidal com uma série formada por uma função de distribuição de probabilidade Normal $\mathcal{N}(0, 0.01^2)$. Em (b), tem-se as magnitudes encontradas pela FT. Em (c), as componentes com magnitudes pouco representativas são zeros removendo assim a influência do componente estocástico. Em (d), a série é imersa de volta no espaço temporal, sem a influência estocástica.

Fonte: Elaborada pelo autor.

xos, é possível observar na Figura 12b uma maior magnitude nas frequências que representam o componente determinístico. Vale mencionar que o gráfico de magnitudes por frequências é simétrico com relação ao ponto central (devido às propriedades da FT (BRACEWELL, 1986)) e, portanto, existem dois picos.

Ao zerar todos os componentes cujas magnitudes não são representativas (Figura 12c) e aplicar a inversa da FT, leva-se os coeficientes restantes de volta ao espaço tempo, construindo a série temporal somente com a influência determinística, como é possível observar na Figura 12d. Nesta abordagem, o componente estocástico pode ser obtido pela subtração do componente determinístico encontrado da série temporal original.

Apesar do bom funcionamento no exemplo anterior, alguns estudos mostram que a alegação de que o componente estocástico está associado aos coeficientes complexos de baixa magnitude não necessariamente é válida (RAMASWAMY, 2006). Ainda, a FT enfrenta problemas ao lidar com dados periódicos e não-estacionários. Além disso, a técnica não localiza frequências com relação ao tempo, mas sim apenas informa quais componentes senoidais somados formam a série original. Por este motivo, a FT impõe um viés senoidal sobre o componente determinístico resultante, mudando o formato do atrator no espaço-fase para uma elipsoide (RIOS *et al.*, 2015; HUANG *et al.*, 1998). Tal mudança faz com que uma próxima observação da série temporal dependa de diferentes valores passados, produzindo, assim, uma topologia matemática anômala e afetando previsões.

3.2 Transformada Wavelet

A transformada Wavelet (do inglês *Wavelet Transform* – WT) (GRAPS, 1995; STOCCHI; MARCHESI, 2018) foi projetada para melhorar os resultados obtidos pela FT, permitindo localizar frequências ao longo do tempo². A WT emprega um conjunto infinito de funções bases $\psi(\cdot)$, denominadas ondaletas (do inglês *wavelets*), que devem atender a propriedades estabelecidas pelas Equações (3.2) e (3.3). Essas propriedades asseguram que a área total (integral) sob a função deve ser igual a zero e sua energia deve ser finita. Tais condições implicam que a *wavelet* tende a oscilar em torno do eixo temporal (primeira propriedade), e que a energia deve estar aglomerada em uma região (segunda propriedade), além disso essas funções devem pertencem ao conjunto $\mathcal{L}^2(\mathbb{R})$ das funções quadrado integráveis.

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad (3.2)$$

$$\int_{-\infty}^{\infty} \psi(t)^2 dt = 1 \quad (3.3)$$

² O termo frequência é tipicamente utilizado dentro do contexto da FT, no contexto da transformada Wavelet é utilizado o termo escala.

Em particular, existe um intervalo finito $[-T, T]$ na reta real em que a energia da ondaleta está concentrada, ou seja, para qualquer $0 \leq \varepsilon \leq 1$ a Equação (3.4) é satisfeita, em que ε é a energia da função fora do intervalo $[-T, T]$. Dessa maneira, uma *wavelet* pode ser compreendida como uma onda com crescimento e decaimento em um período limitado de tempo (PERCIVAL, 2000). Esta característica de concentração de energia das funções *wavelet* é o que propicia a localização da frequência no tempo e a diferencia da FT.

$$\int_{-T}^T \psi(t)^2 dt > 1 - \varepsilon \quad (3.4)$$

Para exemplificar, considere uma *wavelet* formada pela derivada de segunda ordem da função gaussiana $\psi(t) = (1 - 2t^2)e^{-t^2}$ (Figura 13), que concentra toda a sua energia no intervalo $[-4, 4]$, portanto $\varepsilon = \int_{-\infty}^{-4} \psi(t)^2 dt + \int_4^{\infty} \psi(t)^2 dt = 0$.

Figura 13 – Função *wavelet* formada pela derivada de segunda ordem da função gaussiana.

Fonte: Elaborada pelo autor.

No contexto de decomposição de séries temporais, a WT é executada usando diferentes escalas para extrair componentes em diferentes frequências. A primeira escala é resultante de uma execução completa da WT na série temporal para produzir duas sub-séries, cada uma com metade dos elementos do original. Uma metade contém uma aproximação da série original, enquanto a outra contém detalhes tipicamente associados a altas frequências. O mesmo processo pode ser aplicado para extrair componentes adicionais em frequências mais baixas. Pode-se comparar esse processo com um banco de filtros que extrai *features* à partir das frequências mais altas em direção às mais baixas.

A WT mais utilizada é a Haar (HAAR, 1910). Essa se reduz ao cálculo de somas e diferenças entre pares consecutivos de observações da série temporal. Durante o processo de escala, pode-se estudar as frequências, a fim de separar quais componentes são parte do viés estocástico ou determinístico. Mais comumente, estudos referem-se a altas frequências como séries estocásticas enquanto as mais baixas são consideradas determinísticas, o que pode não ser a situação real. Por exemplo, um sistema pode produzir algumas séries em 200Hz,

em que a frequência mais importante para compor o comportamento determinístico está no intervalo 40 e 100Hz enquanto todas as frequências menores que 40Hz podem ser associadas ao componente estocástico. Este é um exemplo real em que os pesquisadores precisam estudar as ondas cerebrais Gamma na tentativa de modelar funções cognitivas saudáveis, como memória e aprendizado (AZCARRAGA; SUAREZ, 2012; RUBIK, 2011).

Embora a WT incorpore recursos adicionais a FT, ela também distorce o espaço-fase, produzindo atratores elipsoidais após a decomposição da série temporal. Isso prejudica a relação temporal entre as observações, influencia fortemente a modelagem e a previsão da série em estudo. A topologia matemática do componente determinístico esperado é novamente alterada, formando o que se espera de uma série sinusoidal (elipsoides com um único buraco) (MENDELSON, 2012). Qualquer mudança indesejável no espaço-fase reflete o viés imposto pela técnica de decomposições. Consequentemente, deve-se encontrar técnicas capazes de fornecer espaços-fases determinísticos esperados (DUARTE *et al.*, 2018).

3.3 Média Móvel

Outro ferramental que tem sido amplamente adotado para decompor séries temporais em componentes estocásticos e determinísticos se dá pela técnica de Média Móvel (do inglês *Moving Average* – MA) (BEVERIDGE; NELSON, 1981; SAID; DICKEY, 1984; GOMEZ; MARAVALL, 1996; ARADHYE; RAO; MOHAMMED, 2018). MA substitui cada observação da série temporal pela média das h observações mais próximas de acordo com o tempo e, por este motivo, é também referida na literatura como um filtro convolucional unidimensional (SAVITZKY; GOLAY, 1964). Esse comportamento tende a remover o componente de maior frequência e que tipicamente está associado ao comportamento estocástico. Por esse motivo, MA pode ser vista como um filtro passa baixa³.

Suponha uma série temporal X , formada por observações $\{x_t, t = 1, \dots, n\}$. A decomposição utilizando MA irá produzir, ao fim, a série temporal Y formada pelas observações $\{y_t, t = h/2, \dots, n - h/2\}$, ou seja, ao se aplicar a decomposição perde-se a informação contida nas $h/2$ observações iniciais e finais. Assim, a série temporal resultante terá o tamanho $n - h$ e pode ser calculada por meio da Equação 3.5.

$$y_t = \frac{1}{h} \left(\sum_{i=0}^{h/2} x_{t-i} + \sum_{i=1}^{h/2} x_{t+i} \right) \quad (3.5)$$

A série temporal final Y é atribuída ao comportamento determinístico e sua subtração da série original resultará no componente estocástico. Para exemplificar todo o processo e o resultado final, considere uma série temporal gerada a partir de uma função senoidal somada

³ Filtro passa-baixa é o nome dado a filtros que atenuam ou removem todos os componentes de altas frequências e que, tipicamente, estão associados a erros de sensores e medições.

a um ruído formado por uma função de distribuição de probabilidades Normal com média zero e variância $0,2^2$. A série temporal resultante foi então submetida à decomposição MA (Figura 14) com janelas de tamanhos diferentes: (a) mostra o efeito da decomposição ao utilizar uma janela $h = 1$ (A Figura 14a) e (b) o efeito ao utilizar uma janela $h = 100$ (Figura 14b). É possível observar que quanto maior o tamanho da janela mais suave e menor será o componente determinístico final.

(a) Resultado atingido com janela de tamanho $h = 1$.

(b) Resultado atingido com janela de tamanho $h = 100$.

Figura 14 – Decomposição da série temporal utilizando médias móveis. A linha vermelha mostra a série inicial e, a preta, a série resultante do processo de médias móveis.

Fonte: Elaborada pelo autor.

Na literatura, esta metodologia é comumente chamada de operadores recursivos (BOX; JENKINS, 1976) ou convolução unidimensional (LI *et al.*, 2019). Vale ressaltar que esse tipo de técnica foi especialmente projetada para suavizar as séries lineares (SAVITZKY; GOLAY, 1964).

3.4 Análise Espectral Singular

A Análise Espectral Singular (do inglês *Singular Spectrum Analysis* – SSA) é uma abordagem livre de parâmetros utilizada para a decomposição de séries temporais arbitrárias em componentes linearmente independentes e interpretáveis. SSA é robusta a variações de tendência da série e, por este motivo, tem sido empregada em cenários reais, nos quais ocorrem mudanças estruturais ao longo do tempo. Apesar disso, essa abordagem parte da premissa de que a série em estudo é linear e estacionária, além de que seu ruído (ou resíduo) segue uma distribuição de probabilidades Normal.

Em linhas gerais, a SSA é dividida em dois estágios: decomposição e reconstrução. No estágio de decomposição, a série temporal é imersa em um espaço vetorial K dimensional e, posteriormente, analisada utilizando a técnica de Decomposição em Valores Singulares (do inglês *Singular Value Decomposition* (SVD)). No estágio de reconstrução, agrupam-se os componentes de acordo com suas respectivas relevâncias, a fim de eliminar ruídos (considerados, neste contexto, como componentes estocásticos).

3.4.1 Estágio de decomposição

O estágio de decomposição da SSA é dividido em duas etapas: imersão e decomposição. Em um primeiro momento, a série temporal é imersa em um espaço vetorial multidimensional. Para exemplificar, considere a série temporal $X = \{x_0, \dots, x_n\}$, em que n é a quantidade de observações. Seja uma janela de tempo de comprimento h , tipicamente definida pelo usuário, com tamanho definido por $2 \leq h \leq n/2$ (GOLYANDINA, 2001), a nova série multidimensional Y imersa em $\mathbb{R}^K$, em que $K = n - h + 1$, é formada pelas subséries $Y = \{Y_1, \dots, Y_K\}$ em que $Y_i = \{x_i, x_{i+1}, \dots, x_{i+h-1}\}^T$ para $1 \leq i \leq K$. Dessa maneira, Y é descrita na forma:

$$Y_n = (Y_1, Y_2, \dots, Y_K) = \begin{pmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_h & x_{h+1} & \dots & x_n \end{pmatrix},$$

a qual forma uma matriz de Hankel (PAN, 2012), i.e., todos os elementos $Y_{i,j}$ para $i + j = c$, em que c é uma constante, são iguais. Para ilustrar, $Y_{1,4} = Y_{4,1} = Y_{2,3} = Y_{3,2} = x_4$, pois em todos os cenários a soma de $i + j$ é igual à constante 5.

A série Y é então multiplicada pela sua transposta, a fim de se obter a matriz de covariância $S = YY^T$. Em seguida, a técnica SVD é aplicada sobre S , decompondo-a como segue $S = U\Sigma V^*$, em que U é uma matriz unitária, Σ é uma matriz diagonal e V^* é a matriz conjugada transposta de V , dado que todas têm as mesmas dimensões $L \times L$ para este problema específico (GOLUB; KAHAN, 1965) (por isso também é possível empregar a Análise de Componentes Principais – do inglês *Principal Component Analysis* (PCA) (STRANG, 2003)).

A fim de interpretar o estágio de decomposição, note que a SVD produz uma base em um espaço $\mathcal{R}^h$, que permite analisar a relevância de cada uma dessas h dimensões. Uma vez que S é resultante da multiplicação YY^T , as covariâncias obtidas representam relações entre os vetores do espaço linha de Y , tal qual realizado por técnicas de *Subspace learning* (SL) (LU; PLATANOTIS; VENETSANOPOULOS, 2013). Maiores valores de covariância implicam em maior similaridade entre vetores linha, os quais correspondem a subséries no domínio do tempo. Logo, a relevância dos autovalores associados aos autovetores da base encontrada para o espaço $\mathbb{R}^h$ está intimamente ligada ao número de subséries mais similares, ou seja, que ocorrem com maior frequência no domínio do tempo. Ao se selecionar os autovalores mais significativos dessa

base, pode-se reconstruir os formatos (no domínio do tempo) ou subséries que constituem os componentes mais comuns da série original.

3.4.2 Estágio de reconstrução

Cada par de autovalores produzido pela SVD (matriz diagonal Σ) está relacionado a um componente senoidal, sendo um deles referente à dimensão real e o outro à imaginária. Logo, para reconstruir a série temporal com os 3 componentes senoidais mais relevantes, deve-se considerar os 6 primeiros autovalores para indicar os autovetores relevantes nesse estágio. Obviamente, se os autovalores menos significativos forem selecionados, a série reconstruída conterá apenas componentes menos relevantes e que estão associados aos componentes de menor amplitude de sinal (os quais, nessa situação, são, por definição, associados aos ruídos).

Dessa maneira, sejam os autovalores obtidos para a SVD $\lambda = \{\lambda_1, \dots, \lambda_h\}$. Têm-se que para reconstruir um componente senoidal de relevância $j = 1, \dots, \frac{h}{2}$ (cujo menor valor indica maior relevância), deve-se considerar os autovalores $\lambda_{2 \times j - 1}$ e $\lambda_{2 \times j}$. Após definir o conjunto de componentes senoidais $G = \{2 \times j - 1, 2 \times j, \dots\}$, para todo j definido pelo usuário, deve-se utilizar da matriz U produzida pela SVD, a qual corresponde aos autovetores que aplicam uma transformação linear sobre a matriz S para produzir um novo espaço $\mathbb{R}^h$, e resolver $\tilde{Y} = \sum_{g \in G} U_{g,k} U_{g,k}^T Y$.

A resultante $\tilde{Y}$ é uma matriz de Hankel que aproxima Y . Logo, se todos os possíveis senoidais forem considerados, i.e., $G = \{1; 2; \dots; h\}$, $\tilde{Y} = Y$. Devido ao fato de $\tilde{Y}$ ser uma matriz de Hankel, ela pode ser reconstruída no domínio temporal, para cada instante t , a partir da média dos elementos presentes em cada uma de suas diagonais cuja soma de índices linha l e coluna c seja igual a t .

3.5 Empirical mode decomposition

Ao contrário das outras abordagens mencionadas, a Decomposição de Modo Empírico (do inglês *Empirical Mode Decomposition* – EMD) manipula séries temporais independentemente da linearidade, estocasticidade e estacionariedade presentes (HUANG *et al.*, 1998). Essa técnica decompõe, no domínio do tempo, uma série em diversos monocomponentes denominados Funções de Modo Intrínseco (do inglês *Intrinsic Mode Functions* (IMFs)), permitindo o estudo de suas frequências e amplitudes instantâneas por meio da Análise Espectral de Hilbert (do inglês *Hilbert Spectral Analysis* (HSA)) (HUANG *et al.*, 1998).

O processo realizado pela EMD é ilustrado na Figura 15, na qual, inicialmente, são analisadas todas as observações da série temporal x_t , identificando seus mínimos e máximos locais (Figura 15a). Em seguida, os valores máximos são interligados utilizando uma função *spline* cúbica, formando, assim, um traçado limitante superior u_t (Figura 15b). De maneira

análoga, os mínimos locais são interligados, utilizando também uma *spline* cúbica, formando assim um traçado limitante inferior l_t (Figura 15c).

As funções u_t e l_t formam um envelope para a série temporal, o qual é utilizado para calcular a *spline* média m_t , de acordo com a Equação 3.6. O resultado é uma *spline* cúbica contida dentro do envelope delimitador (Figura 15d).

$$m_t = \frac{u_t + l_t}{2} \quad (3.6)$$

A nova série m_t é subtraída da série temporal original x_t , produzindo assim o primeiro monocomponente candidato de acordo com a Equação 3.7, em que o índice n representa a IMF e k indexa o monocomponente candidato.

$$h_{n,k,t} = x_t - m_t \quad (3.7)$$

Por fim, substitui-se a série original pelo monocomponente candidato encontrado pela Equação 3.7 e, iterativamente, realiza-se todo o processo de encontrar mínimos e máximos locais bem como a definição de um novo monocomponente candidato. Este processo se repete até que, para todo t , m_t é igual a zero ou a diferença entre o número de pontos extremos e o número de vezes que a série cruza o eixo das abscissas (eixo que descreve o domínio temporal) é menor ou igual um. O último monocomponente candidato produzido define, então, a n -ésima IMF, a qual é removida da série original x_t , para que todo o processo seja repetido a fim de encontrar outras IMFs. O processo de encontrar novas IMFs ocorre até que a última delas seja uma função monotônica, impossibilitando, assim, a sua decomposição em uma nova IMF. O último componente é denominado resíduo r_t e é resultante da série temporal original subtraída de todas as IMFs encontradas. Assim, de acordo com o método EMD, a série temporal é formada por k monocomponentes IMF somados ao resíduo final, tal como apresentado na Equação 3.8.

$$x_t = \sum_{n=1}^k h_{n,t} + r_t \quad (3.8)$$

Para exemplificar o processo, considere uma série temporal formada pela soma de um componente determinístico senoidal e outro estocástico, cujas observações foram produzidas por uma distribuição de probabilidades Normal $\mathcal{N}(0; 0,01^2)$, i.e., com média zero e variância $0,01^2$ (Figura 16a). Ao aplicar a técnica EMD, produz-se sete IMFs e um componente de resíduo (Figura 16b). É possível observar que as primeiras IMFs apresentam altas frequências, as quais são tipicamente associadas ao comportamento estocástico. Nem sempre essa associação de estocasticidade com as altas frequências é verificada na prática, porém foi utilizada em diversos cenários estudados por Rios (2013).

Enquanto isso, as IMFs finais apresentam comportamento mais regular e de menor frequência, os quais são tipicamente associados ao componente determinístico. Essas caracterís-

(a) Análise de toda a série temporal com o intuito de identificar os máximos (pontos em preto) e mínimos (pontos em branco) locais.

(b) Os valores máximos são interligados, utilizando uma função *spline* cúbica, formando assim um traçado limitante superior u_t .

(c) Os mínimos locais são interligados, utilizando uma função *spline* cúbica, formando assim um traçado limitante inferior l_t .

(d) A função *spline* média m_t é calculada como a média aritmética dos traçados limitantes, de acordo com a Equação 3.6. O resultado é uma *spline* cúbica contida no intervalo delimitado.

Figura 15 – Etapas envolvidas na decomposição realizada pelo método EMD. Em (a), tem-se os máximos e mínimos locais; em (b), a *spline* sobre os pontos máximos; em (c), a *spline* sobre os pontos mínimos; e, em (d), tem-se a *spline* média.

Fonte: Adaptada de Rios (2013).

(a) Série temporal formada pela soma de um componente determinístico senoidal e um componente estocástico cujas observações são definidas por uma distribuição de probabilidades Normal $\mathcal{N}(0; 0,01)$.

(b) Oito monocomponentes formados após a decomposição da série (a) por meio do emprego da técnica EMD.

Figura 16 – Monocomponentes produzidos pela decomposição de uma série temporal, utilizando a técnica EMD. Em (a), a série temporal formada pela soma de um componente determinístico senoidal e um componente estocástico $\mathcal{N}(0; 0,01^2)$, i.e., que segue uma distribuição de probabilidades Normal com média zero e variância $0,01^2$. Em (b), os oito monocomponentes gerados pela técnica EMD.

Fonte: Elaborada pelo autor.

ticas motivaram [Rios \(2013\)](#) a projetar duas abordagens que analisam as IMFs a fim de produzir os componentes estocástico e determinístico. Essas abordagens são apresentadas a seguir.

3.5.1 Empirical mode decomposition using recurrence plot

A abordagem *Empirical mode decomposition using recurrence plot* (EMD-RP) utiliza a EMD juntamente com a técnica *Recurrence Plot* (RP) ([ECKMANN; KAMPHORST; RUELLE, 1987](#)) para auxiliar na decomposição de uma série temporal em seus respectivos componentes determinístico e estocástico. Cada monocomponente encontrado pela EMD é analisado por meio da técnica RP, a qual aplica o Teorema de Imersão de [Takens \(1981\)](#) a fim de avaliar as dependências entre as observações da IMF quando reconstruídas no espaço-fase. Essa análise utiliza como base uma matriz bidimensional, denominada Matriz de Recorrência, a qual tem células com valor igual a 1 para estados que se mantêm-se vizinhos ao longo do tempo (zero, caso contrário), o que permite analisar recorrências das trajetórias formadas pela IMF. Uma visualização dessa matriz é construída para encontrar pontos isolados que indicam estados que não evoluem em conjunto ao longo do tempo e, por consequência, apontam que a série em questão é predominantemente estocástica. De modo análogo, linhas diagonais correspondem a influências determinísticas.

Essa análise visual é difícil, mesmo para usuários experientes. Assim, como forma de automatizar o processo, é calculada a frequência de linhas de comprimento l presentes no RP. Essa frequência é, então, comparada a um valor limitante que permite determinar se a IMF em estudo é ou não determinística. A soma de todas as IMFs definidas como determinísticas formam o componente determinístico e, a soma das demais IMFs forma o componente estocástico.

Apesar dos bons resultados apresentados, a EMD-RP é sensível ao valor limitante utilizado para definir quais monocomponentes são determinísticos, bem como das dimensões embutida e de separação utilizadas pelo Teorema de Imersão de [Takens \(1981\)](#). Nesse contexto, surgiu a *Empirical mode decomposition using mutual information* (EMD-MI).

3.5.2 Empirical mode decomposition using mutual information

A *Empirical mode decomposition using mutual information* (EMD-MI) é uma abordagem sem parametrização que combina o método EMD com a técnica de Informação Mútua (do inglês *Mutual Information* – MI) ([LI, 1990](#)) para decompor uma série temporal em seus componentes estocástico e determinístico. Essa abordagem, primeiramente, utiliza a EMD para produzir um conjunto de monocomponentes IMF $h_{n,t}$. Em seguida, aplica-se a Transformada de Fourier discreta $F(\cdot)$ sobre cada IMF $h_{n,t}$, a fim de obter os coeficientes complexos $C_{n,t}$ (Equação 3.9).

Uma vez obtidos esses coeficientes, calcula-se o arco-tangente da razão entre a parte imaginária e a real, o que permite obter as fases de cada IMF (Equação 3.10). Essas fases são, então, analisadas com o intuito de verificar o nível de determinismo ou estocasticidade presente

em tais dados.

$$C_{n,t} = F(h_{n,t}) \quad (3.9)$$

$$\theta_{n,t} = \arctan\left(\frac{\Im(C_{n,t})}{\Re(C_{n,t})}\right) \quad (3.10)$$

Considere por exemplo a série temporal formada por um componente determinístico senoidal e um componente estocástico cujas observações foram produzidas por uma distribuição de probabilidades Normal $\mathcal{N}(0;0,1^2)$ (Figura 16a). Neste caso, a EMD produziu sete IMFs e um resíduo, os quais devem ainda ser analisados utilizando a transformada de Fourier. Para cada IMF neste novo espaço, as fases dos coeficientes complexos são encontradas, como pode ser visto na Figura 17a. É possível notar uma nítida diferença entre as fases das primeiras IMFs, que apresentam grandes variações, versus as últimas, que apresentam um casamento de fases mais adequado, o qual é tido como comportamento determinístico, segundo a abordagem EMD-MI.

Para determinar o ponto de separação entre quais IMFs representam o componente estocástico e quais correspondem ao componente determinístico, a técnica MI é aplicada sobre as fases de IMFs consecutivas para estimar a quantidade de informação compartilhada. Rios (2013) observou que, para dados com misturas de componentes estocásticos e determinísticos, as primeiras IMFs apresentam fases discordantes e, em contrapartida, as últimas compartilham muita informação, o que tem relação direta com as influências determinísticas. Surge, assim, um ponto de separação entre IMFs que compõe o componente estocástico (as primeiras) e o determinístico (as últimas ou seguintes).

No exemplo anterior, ao calcular a informação mútua entre as IMFs consecutivas, é possível notar um ponto de separação no gráfico da Figura 17b. O valor obtido para a técnica de MI, quando aplicada sobre as primeiras IMFs, é bem próximo a zero, o que é utilizado por Rios (2013) para caracterizar estocasticidade, uma vez que a aleatoriedade entre fases indica que os coeficientes complexos de Fourier são mal comportados ou aleatórios para aquele monocomponente. Em contrapartida, o valor de MI para as IMFs seguintes é significativamente maior, logo fases transicionam de maneira mais suave para tais monocomponentes, o que é utilizado como critério para definí-los como mais determinísticos. Assim, o ponto de separação é determinado pela mudança abrupta nos valores de informação mútua calculados.

No entanto, apesar da propositura de Rios (2013) ter apresentado excelentes resultados, suas duas abordagens, i.e., EMD-RP e EMD-MI, são limitadas ao viés senoidal representado pelo método EMD. Qualquer outro viés é então reduzido ao senoidal, o que descaracteriza uma vasta gama de séries temporais como as provenientes do Mapa Logístico, atrator do Sistema de Lorenz e de Rössler. Esse viés senoidal deve-se à utilização de funções *spline* cúbicas para a decomposição de séries temporais.

(a) Magnitudes das IMFs no espaço de frequências após transformada de Fourier.

(b) Informação mútua calculada entre pares de IMFs consecutivas. A linha vermelha define a média de todos os valores e é utilizada como critério para separar as IMFs estocásticas das determinísticas.

Figura 17 – Monocomponentes gerados pela decomposição da série temporal utilizando a técnica EMD. Em (a), a série temporal formada pela soma de um componente determinístico senoidal e um componente estocástico cujas observações foram produzidas por uma distribuição de probabilidades Normal $\mathcal{N}(0; 0, 1^2)$. Em (b), os oito monocomponentes gerados pela técnica EMD.

Fonte: Elaborada pelo autor.

3.6 Métodos de decomposição não linear

Em comum, todas as abordagens supracitadas impõem sobre o componente determinístico resultante um viés senoidal. Essa característica modifica o formato do atrator no espaço-fase que tem sua topologia aproximada a uma elipse. A consequência dessa alteração está no aumento do erro em modelos preditivos. Para contornar tais problemas, as técnicas de decomposição não linear Lazy (SCHREIBER, 1993) e GHKSS (GRASSBERGER *et al.*, 1993) trabalham sobre o espaço-fase a fim de manter a topologia e diminuir tal erro.

Utilizando o teorema proposto por Takens (1981) para imergir a série temporal no espaço-fase, a técnica Lazy altera a posição de cada estado do atrator pela posição média dos seus vizinhos. Para isso, o usuário deve determinar um raio ε que forma, para cada estado s_i no espaço-fase, uma vizinhança $\mathcal{U}_\varepsilon^{(n)}$ que contém os estados vizinhos $\{s_{i^*}\}$ tal que $\|s_i - s_{i^*}\| < \varepsilon$. Este raio que, determina a vizinhança, deve ser grande o suficiente para conter toda dispersão provocada pela adição do ruído (componente estocástico), mas ainda menor que o raio da curvatura natural do atrator. Todo processo é repetido iterativamente até que o número máximo de iterações, estabelecido pelo usuário, aconteça ou não haja mais movimentações dos estados em relação a iteração anterior. O atrator final é associado ao componente determinístico e sua subtração da série original ao componente estocástico. Na prática, o algoritmo contrai todos os pontos em direção ao esqueleto⁴ do atrator, produzindo, ao fim, uma versão simplificada e sem ruído.

Entretanto, as condições impostas sobre o raio, em alguns casos, pode não ser factível, deixando a cargo do usuário escolher diversos valores para o raio e avaliar o resultado final. Ainda, a depender do ruído e raio escolhido, a vizinhança pode conter apenas o próprio ponto, o que implica que nenhuma modificação ou filtragem é realizada. Por fim, dada a metodologia, os primeiros e últimos $(m - 1)/2$ observações da série temporal, em que m é a dimensão de embutir, não sofrem modificações. Isso acontece pois ao imergir o atrator de volta no espaço tempo apenas os valores da dimensão do meio são consideradas.

Outro algoritmo que também trabalha sobre o espaço-fase, o GHKSS (GRASSBERGER *et al.*, 1993) assume a hipótese de que a série temporal em estudo foi produzida por um sistema dinâmico de baixa dimensionalidade somado a um ruído aleatório de alta dimensionalidade. Isso implica que, em um espaço de alta dimensão arbitrário, a parte determinística dos dados residiria em uma variedade matemática de dimensão menor e o ruído seria responsável por espalhar os estados nessa variedade.

Suponha uma variedade q -dimensional $\mathcal{M}$ que contenha todas as trajetórias do atrator formado por um sistema dinâmico (Palis Junior, 1977). De acordo com o teorema de imersão, em uma dimensão suficientemente grande, existe um único atrator que representa o comportamento

⁴ Esqueletos são descritores de objetos que representam uma versão simplificada da geometria e topologia de uma estrutura n -dimensional (TAGLIASACCHI; ZHANG; COHEN-OR, 2009).

desse sistema. Devido à presença do ruído, esta condição não mais é satisfeita (HEGGER; KANTZ; SCHREIBER, 1999). Neste contexto, a técnica GHKSS tenta calcular para cada estado s_i qual a correção Θ_i , tal que $s_i - \Theta_i \in \mathcal{M}$, e Θ_i seja ortogonal a $\mathcal{M}$. Em suma, a técnica GHKSS transforma a filtragem em um problema de otimização para encontrar todos os Θ_i .

3.7 Considerações Finais

Há a necessidade de decompor séries temporais em seus respectivos componentes, determinístico e estocástico, para que cada um seja tratado com as ferramentas mais adequadas ao seu viés. Assim, as técnicas descritas neste capítulo são comumente empregadas para esse tipo de tarefa. Em comum, as técnicas FT, WT, MA, SSA, EMD-RP e EMD-MI, ao final, produzem componentes determinísticos com forte viés senoidal. Essa característica pode ser verificada no atrator do componente resultante que, no espaço-fase, apresenta formato elipsoidal, tipicamente diferente do formato original. Essa distorção prejudica a análise de recorrências e tendências, o que produz efeitos negativos para o algoritmo de modelagem e predição de séries temporais. Por sua vez, as técnicas de decomposição e filtragem não linear Lazy e GHKSS apresentam alta complexidade, sensibilidade aos hiper-parâmetros e deixam a cargo de leitor sua parametrização. Assim, surge a necessidade de uma técnica que seja capaz de decompor uma série temporal em suas influências, preservando, contudo, o viés original do componente determinístico e, consequentemente, o formato do atrator no espaço-fase, conforme será abordado no próximo capítulo.

ABORDAGENS PROPOSTAS PARA DECOMPOSIÇÃO DE SÉRIES TEMPORAIS

As principais técnicas de decomposição que compõem o estado da arte impõem um viés senoidal sobre o componente determinístico das séries temporais. Essa imposição pode destruir as relações entre as amostras da série temporal, distorcendo recorrências e tendências e, dessa forma, prejudicando o processo de predição e análise. A consequência desta distorção pode ser observada no espaço-fase onde o formato do atrator é desfigurado. Há, assim, a necessidade de uma técnica capaz de decompor uma série temporal em seus respectivos componentes estocástico e determinístico, preservando, contudo, o viés determinístico original da série. Para isto, o processo de decomposição deverá ser realizado sobre o espaço-fase, onde as relações, tendências e recorrências estão desdobradas. O uso desse espaço assegura que o viés determinístico da série seja preservado durante o processo de decomposição tal como explorado pelas técnicas Lazy ([SCHREIBER, 1993](#)) e GHKSS ([GRASSBERGER et al., 1993](#)).

Para ilustrar, considere a série temporal apresentada na Figura 18c, a qual é resultante da soma termo-a-termo de um componente determinístico produzido pelo Sistema de Lorenz com $\sigma = 10$, $\rho = 21$ e $\beta = 8/3$ (Figura 18a), e outro componente estocástico (Figura 18b) produzido por uma distribuição de probabilidades Normal com média zero e variância 0,25, i.e., $\mathcal{N}(0;0,25)$. Ao imergir a série puramente determinística (Figura 18a) no espaço-fase (utilizando o Teorema de Imersão proposto por [Takens \(1981\)](#) – Seção 2.3.2), as relações de dependência entre suas observações são desdobradas, formando um atrator característico que remete ao símbolo do infinito (Figura 19a). A adição do componente estocástico distorce as relações entre os estados no espaço-fase, preservando, contudo, características topológicas do atrator (Figura 19b) como o seu formato, número de buracos e, em menor grau, a vizinhança entre estados ([MENDELSON, 2012](#)). Essa propriedade pode ser observada na Figura 19b, que é resultante da imersão da série temporal formada pela soma dos componentes determinístico e estocástico (Figura 18c).

(a) Componente determinístico produzido pelo Sistema de Lorenz.

(b) Componente estocástico produzido por uma distribuição de probabilidades Normal $\mathcal{N}(0, 0.25)$.

(c) Soma termo-a-termo dos comportamentos determinístico e estocástico.

Figura 18 – A série temporal (c) é produzida a partir da soma dos componentes determinístico (a) e estocástico (b).

Fonte: Elaborada pelo autor.

É importante ressaltar que caso o componente estocástico seja predominante, o atrator no espaço-fase não remeterá ao formato do atrator do componente determinístico. Assim, o alvo do estudo descrito neste documento são séries em que o componente estocástico é adicionado a um componente determinístico que domina a série original e, por esse motivo, supõe-se que seu atrator sempre mantém os mesmos aspectos topológicos do atrator determinístico. Além disto, para a imersão nesse novo espaço, faz-se necessário determinar as dimensões de separação τ e embutida m , tal que ocorra o desdobramento das relações temporais no espaço-fase. Ao subestimar esses valores pode-se produzir atratores pouco representativos e, ao contrário, superestimar tais valores pode-se produzir espaços demasiadamente complexos.

Após imersão no espaço-fase, o processo de decomposição da série temporal é realizado por meio da contração dos estados em direção ao esqueleto do atrator, removendo assim toda a influência estocástica, a qual causa variações indevidas ao componente determinístico. Define-se

Figura 19 – Comparação entre o espaço-fase produzido pelo Sistema de Lorenz (a) e por sua combinação com um componente estocástico formado por uma distribuição de probabilidades Normal $\mathcal{N}(0; 0, 25)$ (b).

Fonte: Elaborada pelo autor.

esqueleto como a linha interna ao objeto, que é equidistante às bordas. Esqueletos tipicamente enfatizam propriedades geométricas e topológicas do objeto como buracos, direções, larguras, alturas, entre outras (OGNIEWICZ, 1995; TAGLIASACCHI; ZHANG; COHEN-OR, 2009).

Considere o atrator do exemplo anterior, produzido pela imersão no espaço-fase da série temporal formada pelo Sistema de Lorenz somado a um componente estocástico cujas observações foram produzidas por uma distribuição de probabilidades Normal $\mathcal{N}(0; 0, 25)$. A Figura 20 mostra como deve ser a contração dos pontos em direção ao esqueleto do atrator. Em preto são apresentados os estados imersos e, em vermelho, o esqueleto final que representa o componente determinístico sem qualquer influência estocástica.

Diversas metodologias podem ser empregadas para se encontrar o esqueleto de uma nuvem de pontos (CAO *et al.*, 2010; DEY; SUN, 2006; TAGLIASACCHI; ZHANG; COHEN-OR, 2009), tal qual os estados presentes no espaço-fase podem ser vistos. Essa operação representa a subtração do componente estocástico da série temporal original, contudo ainda presente no espaço-fase. Com os pontos já atraídos, pode-se retornar o atrator final para o espaço tempo, obtendo-se o componente determinístico. Logo, pode-se remover o componente encontrado da série temporal original, produzindo-se, assim, o componente estocástico.

Algumas séries apresentam atratores com trajetórias mais espessas, i.e., as trajetórias não formam exatamente um esqueleto no espaço-fase. Por este motivo a metodologia descrita acima, apesar de eficaz para o cálculo do esqueleto, não leva em consideração tais características que algumas séries temporais apresentam, tais como o Sistema de Lorenz e o de Rössler. Ao

Figura 20 – O processo de decomposição proposto deve convergir todos os estados em direção ao esqueleto do atrator. Em (a), mostra-se o atrator do Sistema de Lorenz com influência estocástica e, em vermelho, seu possível esqueleto. Em (b), uma secção transversal em certa região do atrator é apresentada a fim de ilustrar a visualização do esqueleto.

Fonte: Elaborada pelo autor.

transformar esses atratores em seus respectivos esqueletos, o espaço temporal sofre uma mudança significativa das frequências, além de uma modularização das amplitudes da série. Para contornar esses problemas, deve-se buscar compreender o número de iterações necessárias para obter bons resultados com as técnicas iterativas, de modo a parar todo o processo antes que todos os estados converjam para o esqueleto central.

Com base nesses argumentos, foram desenvolvidas as técnicas **Spring** (DUARTE *et al.*, 2018) (Seção 4.1) e suas variações **Spring*** (Seção 4.3) e **Spring Time Domain** (Seção 4.2). Em comum, todas essas técnicas trabalham sobre o espaço-fase a fim de preservar o viés determinístico e, consequentemente, melhorar os resultados de decomposição e predição.

4.1 Spring

Utilizando o teorema proposto por Takens (1981) (Seção 2.3.2), a técnica **Spring** inicia seu processo imergindo a série temporal no espaço-fase, tal como na Equação 4.1, a fim de compreender e detectar padrões de comportamento e recorrências. Segundo tal equação, x_i é a i -ésima observação da série temporal no espaço tempo, τ é a dimensão de separação, m é a dimensão de embutir, e s_i é o i -ésimo estado no espaço-fase. Cabe ressaltar que a técnica **Spring** parte do pressuposto de que as dimensões de embutir e de separação são conhecidas pelo usuário e recebidas como parâmetros. Uma possível metodologia para estimar tais parâmetros foi discutida nas Seções 2.3.4 e 2.3.3.

$$s_i = (x_i, x_{i+\tau}, \dots, x_{i+(m-1)\tau}) \quad (4.1)$$

Seja $\mathcal{S} = \{s_0; \dots; s_n\}$ o conjunto de estados que representa o atrator da série temporal imersa no espaço-fase. $\mathcal{D}$ é uma matriz de distâncias Euclidianas, tal que cada elemento é dado por $\mathcal{D}_{ij} = d(s_i, s_j)$ para $i = 1, \dots, n$ e $j = 1, \dots, n$, ou seja, a matriz fornece a distância de cada estado s_i em relação aos demais. A partir desta matriz $\mathcal{D}$ é possível calcular os k vizinhos de cada estado s_i . Assim, iterativamente, as posições dos estados são modificadas de acordo com uma média ponderada da posição dos k vizinhos mais próximos (BISHOP, 2006) (Figura 21).

Figura 21 – Em vermelho, ao centro, um estado $s_i \in \mathbb{R}^m$ qualquer e, em preto, os k vizinhos mais próximos, sabendo que a projeção bidimensional pode confundir essa vizinhança que está, de fato, definida em $\mathbb{R}^3$. A distância Euclidiana entre eles é representada pelas linhas em azul.

Diferentemente da técnica Lazy (SCHREIBER, 1993), que utiliza um raio fixo ε para determinar quais vizinhos considerar, a **Spring** utiliza um raio adaptativo a fim de lidar com as diferenças de densidades entre os estados do atrator. Esta flexibilidade soluciona um dos principais problemas da técnica Lazy relacionados ao raio (vide Seção 3.6). Para isso, o usuário deve informar o quantil a ser considerado que, além de ser intuitivo, indica o percentual de estados a ser considerado como vizinho. Como consequência, em regiões de maior densidade, o raio utilizado será menor e, contrariamente, regiões de menor densidade necessitarão de raios maiores. Ainda, dada esta adaptação, não há a possibilidade de nenhum estado ser ignorado no processo de filtragem por não apresentar qualquer vizinho, o que difere da abordagem proposta em Lazy (SCHREIBER, 1993). Em suma, a cada iteração, calcula-se para cada estado s_i seu raio e sua nova posição.

Após determinar os estados vizinhos que serão considerados, a posição do estado s_i é alterada para a média ponderada da posição dos vizinhos. O peso da ponderação é inversamente proporcional à distância, i.e., quanto mais distante um estado vizinho s_j está de s_i , menor será sua influência sobre a nova posição. Este processo é feito para todos os estados $s \in \mathcal{S}$, resultando em uma contração do atrator em direção ao seu esqueleto. Todo este processo é repetido até que o deslocamento¹ médio dos estados seja menor que um ε fornecido como parâmetro pelo usuário,

¹ Deslocamento: distância Euclidiana entre a posição da iteração t e posição da iteração $t - 1$ de um mesmo estado s_i .

ou até atingir um número máximo de iterações. Empiricamente, observou-se que o número ideal de iterações é aproximadamente dois, podendo, eventualmente, chegar até cinco.

Para ilustrar o processo, considere uma série temporal formada por um componente determinístico senoidal somada a um ruído branco que segue uma distribuição de probabilidades Normal. O atrator formado por essa série no espaço-fase remete ao atrator elíptico, característico do componente senoidal, distorcido pela presença do componente estocástico (Figura 22a). Ao aplicar a **Spring** sobre essa configuração inicial, os estados são alterados em direção ao esqueleto do atrator (22b, 22c, 22d, 22e). Este processo é repetido até que o deslocamento médio seja menor que um parâmetro ε determinado pelo usuário, i.e., até que a posição dos elementos na iteração i e $i - 1$ seja praticamente a mesma, como é possível observar nas Figuras 22e e 22f.

Figura 22 – Exemplo de execução do algoritmo **Spring**. Em (a), espaço-fase de uma série temporal formada pela soma de uma função senoidal com um ruído branco, que segue uma distribuição de probabilidades Normal. As iterações do algoritmo são mostradas em (b), (c), (d) e (e). Em (f), é apresentada a última iteração e o estado de convergência do algoritmo.

Fonte: Elaborada pelo autor.

Uma analogia que ilustra e facilita o entendimento do processo é: considere um sistema de molas que conecta cada estado $s_i \in \mathcal{S}$ aos seus k vizinhos mais próximos. A ponderação calculada, utilizando as distâncias entre os estados, será a constante elástica das molas utilizada na Lei de Hooke (TIPLER, 2009), i.e., quanto menor a distância, mais rígida será a mola. A técnica **Spring** simula o processo de convergência desse sistema até que se atinja o equilíbrio físico.

Todo esse processo de decomposição está transcrito no Algoritmo 1 em que: ts é a

série temporal a ser decomposta, n é o número máximo de iterações, q é o quantil utilizado para determinar o tamanho do raio, ε o deslocamento máximo para convergência, m e τ são as dimensões de embutir e de separação, respectivamente.

Algoritmo 1 – Pseudo algoritmo da técnica **Spring**.

função SPRING($ts, n, q, \varepsilon, m, \tau$)

$\mathcal{S} \leftarrow \text{EMBED}(ts, m, \tau)$

$\mathcal{D} \leftarrow \text{DISTMAT}(\mathcal{X})$

para 1 to n **faça**

para todo $s_i \in \mathcal{S}$ **faça**

$r \leftarrow \text{QUANTIL}(s_i, \mathcal{S}, q)$

$\text{nn_dist} \leftarrow \text{NEIGHBORDISTANCE}(x, \text{nn_pos})$

$\text{nn_simi} \leftarrow \text{NEIGHBORSIMILARITY}(\text{nn_dist})$

$s'_i \leftarrow \text{nn_pos} * \text{nn_simi}$

fim para

$\text{disp} \leftarrow \text{AVERAGEDISPLACEMENT}(\mathcal{S}, \mathcal{S}')$

$\mathcal{S} \leftarrow \mathcal{S}'$

se $\text{disp} \leq \varepsilon$ **então**

 break

fim se

fim para

retorna TOTIMESPACE($\mathcal{X}$)

fim função

▷ Imerge a série temporal no espaço-fase

▷ Calcula a matriz de distâncias

▷ Número máximo de iterações

▷ Para cada estado do atrator

▷ Calcula raio baseado no quantil

▷ Distâncias dos vizinhos para $\dot{x}_j$

▷ Similaridade normalizada em $[0, 1]$

▷ Calcula a nova posição do estado

▷ Deslocamento médio dos estados do atrator

▷ Atualiza a posição do atrator

▷ Critério de parada

4.2 Spring Time Domain

Apesar dos bons resultados, alguns sistemas, como o de Lorenz e Rössler, produzem séries temporais cujos atratores apresentam transições suaves e bem comportadas (Figura 23a). O mesmo não é verificado em séries que utilizam mapas como processo gerador determinístico (Figura 23b). Ainda, os atratores formados por sistemas apresentam uma espessura natural e a utilização de técnicas que buscam seu esqueleto podem deteriorar o comportamento da série, prejudicando a etapa de predição. Neste contexto foi proposta a técnica **Spring Time Domain** (Spring-TD).

Levando em consideração as características das transições suaves entre os estados e tentando preservar a espessura natural do atrator, a **Spring-TD** utiliza splines cúbicas (HASTIE; TIBSHIRANI; FRIEDMAN, 2016) para remover a influência estocástica do atrator. Após imergir a série temporal no espaço-fase, a técnica navega no atrator utilizando uma janela W com comprimento L centralizada em s_i , i.e., $W = \{s_{i-L/2-1}, \dots, s_i, \dots, s_{i+L/2-1}\}$. Os estados em W são interpolados por uma função spline cúbica, e a posição do estado central da janela é substituída pela sua projeção nessa função. A janela, por fim, é deslocada em uma unidade, e

Figura 23 – Comparação entre o atrator do Sistema de Lorenz (a) e o atrator do Mapa Logístico (b). É possível observar, em (a), transições suaves entre os estados e, em (b), transições abruptas.

Fonte: Elaborada pelo autor.

o processo é repetido para todo estado s_i pertencente ao espaço-fase $\mathcal{S}$. Este processo produz uma suavização nas transições entre os estados e, conseqüentemente, nas trajetórias do atrator. Todo este processamento é repetido, iterativamente, até que deslocamento médio dos estados seja menor que um ε fornecido pelo usuário.

A Figura 24 ilustra todo o processo de decomposição. Considere a mesma série explorada nos exemplos anteriores, a qual é formada por um componente determinístico senoidal somado a um ruído branco com distribuição de probabilidades Normal. A Figura 24a mostra o posicionamento inicial dos estados do atrator, que tem formato elíptico. O algoritmo **Spring-TD** suaviza as transições entre os estados, removendo, assim, a influência estocástica do atrator (24b, 24c, 24d, 24e). Este processo é repetido até que o critério de parada seja satisfeito (Figura 24f). Resumidamente, todo esse processo de decomposição está transcrito no Algoritmo 2, em que: ts é a série temporal a ser decomposta, n é o número máximo de iterações, l é o comprimento da janela W , ε o deslocamento máximo para convergência, m e τ são as dimensões de embutir e de separação, respectivamente.

A eficácia da **Spring-TD**, no entanto, é observada apenas para atratores cuja transição é suave. Assim, obtém-se melhores resultados ao utilizar a **Spring** tradicional e a **Spring-TD** a depender das características de transição dos estados do atrator. Portanto, faz-se necessária uma metodologia capaz de identificar, automaticamente, quando a transição entre estados é suave. Isso possibilitaria automatizar um processo que decidiria qual técnica seguir no processo de decomposição: **Spring** tradicional ou **Spring-TD**. Para isso, calcula-se a distância média $\bar{d}$ de cada estado do atrator para o próximo estado, i.e., a distância entre o estado s_i e s_{i+1} tal como descrito na Equação 4.2 em que: n é o número de estados do atrator e $d(.,.)$ é a função,

Figura 24 – Exemplo de execução do algoritmo **Spring Time Domain (Spring-TD)**. Em (a), espaço-fase de uma série temporal formada pela soma de uma função senoidal com uma distribuição de probabilidades Normal. As iterações do algoritmo são mostradas em (b), (c), (d) e (e). Em (f), é apresentada a última iteração e o estado de convergência do algoritmo.

Fonte: Elaborada pelo autor.

Algoritmo 2 – Pseudo algoritmo da técnica **Spring Time Domain (Spring-TD)**.

função SPRING TIME DOMAIN($ts, n, l, \varepsilon, m, \tau$)

$\mathcal{S} \leftarrow \text{EMBED}(ts, m, \tau)$

$r \leftarrow \text{MAX}(\mathcal{D})$

para 1 to n **faça**

para todo $s_i \in \mathcal{S}$ **faça**

$\text{nn_pos} \leftarrow \text{WN}(s_i, \mathcal{D}, l)$

$s'_i \leftarrow \text{SMOOTH_SPLINE}(s_i, \text{nn_pos})$

fim para

$\text{disp} \leftarrow \text{AVERAGEDISPLACEMENT}(\mathcal{S}, \mathcal{S}')$

$\mathcal{S} \leftarrow \mathcal{S}'$

se $\text{disp} \leq \varepsilon$ **então**

break

fim se

fim para

retorna TOTIMESPACE($\mathcal{S}$)

fim função

▷ Imerge a série temporal no espaço-fase

▷ Raio inicial igual a máxima distância entre estados

▷ Número máximo de iterações

▷ Para cada estado do atrator

▷ Recupera a posição dos vizinhos dentro do janelas l

▷ Calcula nova posição de s_i baseado em uma projeção deste na spline cúbica

▷ Deslocamento médio dos estados do atrator

▷ Atualiza a posição do atrator

▷ Critério de parada

tipicamente Euclidiana, que mensura a distância entre dois estados consecutivos do atrator.

$$\bar{d} = \frac{1}{n-1} \sum_{i=1}^{n-1} d(s_i, s_{i+1}) \quad (4.2)$$

Empiricamente observou-se que ao analisar uma série temporal normalizada no intervalo $[-1, 1]$, os atratores que apresentam transições suaves entre os estados têm distância média menor que os atratores com transições abruptas (empiricamente, $\bar{d} \leq 0.5$). Desse modo é possível automatizar o processo e obter a melhor decomposição para cada cenário.

4.3 Spring*

O resultado do processo de decomposição pode variar a depender dos hiper-parâmetros fornecidos pelo usuário. Desse modo, apesar das técnicas **Spring** e **Spring Time Domain** serem facilmente compreendidas, ainda cabe ao usuário entender seus detalhes específicos para obter os melhores resultados possíveis. Neste contexto, surge a necessidade de uma técnica que apresente bons resultados no processo de decomposição e não requeira que o usuário defina parâmetros críticos, tal como a EMD-MI (Seção 3.5.2).

Utilizando o mesmo conceito da técnica **Spring**, que contrai os estados do atrator em direção ao esqueleto para decompor a série temporal, a técnica **Spring*** utiliza o número de falsos vizinhos mais próximos para estimar a dimensão de embutir (Seção 2.3.4) e a informação mútua para a dimensão de separação (Seção 2.3.3). Com tais valores, a série temporal em estudo é imersa no espaço-fase, formando o atrator $\mathcal{S} = \{s_0, \dots, s_n\}$. Uma matriz de distâncias Euclidianas $\mathcal{D}$ é calculada, tal que cada elemento é dado por $\mathcal{D}_{ij} = d(s_i, s_j)$ para $i = 1, \dots, n$ e $j = 1, \dots, n$, ou seja, a matriz fornece a distância de cada estado $s_i \in \mathcal{S}$ em relação aos demais.

Essa matriz possibilita calcular quais são os estados vizinhos que devem ser considerados no cálculo da contração do atrator. No contexto da **Spring**, o usuário precisa fornecer o quantil para que o raio fosse adaptativamente estimado, bem como o número de iterações. Diferentemente, a **Spring*** utiliza o teste de hipótese Ljung-Box (LJUNG; BOX, 1978) para estimar tais parâmetros. Esse teste permite mensurar se uma dada série temporal $X = \{x_t, t = 1, \dots, n\}$ é formada exclusivamente por observações aleatórias. Para isso, sua hipótese nula H_0 supõe que os dados sejam independentes, i.e., as correlações entre as observações x_t da série temporal em estudo são zero. Assim, qualquer correlação observada nos dados resulta da aleatoriedade do processo que, em neste contexto, está associada ao componente estocástico.

O processo de decomposição da **Spring*** começa com o componente estocástico nulo e o raio sendo igual à máxima distância entre os estados do atrator. Em uma primeira iteração, a posição de cada estado s_i é alterada para a média ponderada da posição dos respectivos vizinhos contidos dentro deste raio. Esse novo atrator é imerso de volta no espaço tempo e sua diferença para a série temporal inicial é somada ao componente estocástico que é validado com o teste

Ljung-Box. Caso o p -value, dado pelo teste, seja menor que 0,05 a decomposição calculada nessa iteração de fato removeu influências estocásticas, logo mais uma iteração é realizada com o novo atrator, mantendo-se o raio. Caso contrário, a modificação nos estados do atrator é desconsiderada, o raio é diminuído a sua metade e o processo de decomposição no espaço-fase é então repetido.

Ao decorrer do processo e da diminuição do raio, esse chegará a um tamanho que não mais conterá estado dentro da vizinhança de cada s_i . Nesse momento, como consequência, a posição de cada estado s_i não será alterada, garantindo a convergência do processo e finalização da decomposição. Esse atrator final é imerso de volta no espaço-tempo formando o componente determinístico e sua subtração da série temporal inicial fornece o componente estocástico.

Cabe aqui ressaltar que o teste de hipótese Ljung-Box necessita do valor de um lag h que, tipicamente, deve ser fornecido pelo usuário. Para determinar esse número e evitar que o processo de decomposição volte a necessitar de um hiper-parâmetro, utilizou-se o valor proposto por Burns (2003). Naquele estudo foram realizados diversos experimentos, concluindo-se que o teste Ljung-Box apresenta melhor resultados com $h \leq 0,05n$, ou seja, 5% do comprimento da série em estudo. Esse valor é então utilizado em todo o processo de decomposição.

Todo esse processo é ilustrado na Figura 25. Dada uma série temporal construída a partir de um componente determinístico senoidal somado a um ruído branco, que segue uma distribuição de probabilidades Normal, seu atrator terá formato elíptico distorcido pela presença do componente estocástico (Figura 25a). Ao aplicar a técnica **Spring*** sobre essa configuração inicial, os estados são alterados em direção ao esqueleto do atrator (25b, 25c, 25d, 25e). Esse processo é repetido até que o raio utilizado seja pequeno o suficiente para alterar a posição de cada estado. Devido à metodologia aplicada pela **Spring***, todo esse processo requer mais iterações que sua versão parametrizável como neste exemplo com 14 iterações até a convergência (Figura 25f).

Para facilitar a compreensão e possibilitar a replicação da metodologia, todo processo está transcrito no Algoritmo 3, em que ts é a série temporal a ser decomposta.

4.4 Considerações Finais

A maioria das técnicas de decomposição de séries temporais impõe um viés senoidal sobre o componente determinístico resultante. Esse viés pode prejudicar a análise e modelos de predição. Ademais, técnicas que se propuseram a manter o viés apresentam difícil parametrização e alguns problemas relacionados à densidade de pontos no espaço-fase. Neste contexto, surgem as técnicas **Spring**, **Spring*** e **Spring Time Domain**. Em comum, todas elas se propõem a decompor séries temporais preservando o viés determinístico e diminuindo a complexidade de utilização. Cada uma das variações da **Spring** tem características próprias e deve ser utilizada criteriosamente, a depender do cenário alvo. A **Spring** tradicional é mais complexa de se

Figura 25 – Exemplo de execução do algoritmo **Spring***. Em (a), espaço-fase de uma série temporal formada pela soma de uma função senoidal com uma distribuição de probabilidades Normal. As iterações do algoritmo são mostradas em (b), (c), (d) e (e). Em (f), é apresentada a última iteração e o estado de convergência do algoritmo.

Fonte: Elaborada pelo autor.

parametrizar, porém apresenta maior desempenho quando há uma transição abrupta entre os estados do atrator. Por sua vez, caso a transição seja suave, a **Spring Time Domain** apresenta a mesma complexidade de parametrização, mas com resultados melhores. Em um cenário em que o usuário não tenha conhecimento sobre o fenômeno estudado, nem sobre as características das técnicas de decomposição, a **Spring*** proporciona uma decomposição, ainda que sub-ótima, para as séries em estudo. Para garantir a qualidade dessas novas abordagens, elas foram submetidas a um extenso processo de avaliação e comparadas às técnicas de decomposição do estado-da-arte.

Algoritmo 3 – Pseudo algoritmo da técnica **Spring***.**função** SPRING*(ts) $\tau \leftarrow \text{EST_DIM_SEP}(ts)$ $m \leftarrow \text{EST_DIM_EMB}(ts)$ $\mathcal{S} \leftarrow \text{EMBED}(ts, m, \tau)$ $\mathcal{D} \leftarrow \text{DISTMAT}(\mathcal{X})$ $r \leftarrow \text{MAX}(\mathcal{D})$ **enquanto** 1 **faça****para todo** $s_i \in \mathcal{S}$ **faça** $\text{nn_pos} \leftarrow \text{NNR}(s_i, \mathcal{D}, r)$ $s'_i \leftarrow \text{NOVA_POS}(s_i, \text{nn_pos})$ **fim para****se** $\mathcal{S} - \mathcal{S}' = 0$ **então**

break

fim se $\text{nseries} \leftarrow \text{TOTIMESPACE}(\mathcal{S}')$ $\text{pv_lj} \leftarrow \text{LJUNG-BOX}(ts - \text{nseries})$ **se** $\text{pv_lj} > 0.05$ **então** $\mathcal{S} = \mathcal{S}'$ **senão** $r = r * 0.5$ **fim se****fim enquanto****retorna** $\text{TOTIMESPACE}(\mathcal{S})$ **fim função**

- ▷ Estima a dimensão de separação utilizando informação mútua (Seção 2.3.3)
- ▷ Estima a dimensão de embutir utilizando Falsos Vizinhos mais próximos (Seção 2.3.4)
- ▷ Imerge a série temporal no espaço-fase
- ▷ Calcula a matriz de distância
- ▷ Raio inicial igual a máxima distância entre estados
- ▷ Recursão de cauda transformado em laço
- ▷ Para cada estado do atrator
- ▷ Recupera a posição dos vizinhos dentro do raio r
- ▷ Calcula nova posição de s_i baseado em sua posição e dos vizinhos
- ▷ Critério de parada por não haver modificação
- ▷ Imerge atrator final de volta para o espaço tempo
- ▷ p -value do teste de Ljung-Box
- ▷ Aceita hipótese nula. Nova iteração sem diminuir o raio r
- ▷ Rejeita hipótese nula. Ignora qualquer modificação no atrator e refaz decomposição com raio menor

MEDIDAS DE AVALIAÇÃO DE TÉCNICAS DE DECOMPOSIÇÃO

Na literatura, diversas medidas são comumente adotadas para comparar e determinar semelhanças entre séries temporais. Essas são tipicamente adotadas para avaliar modelos preditivos em que as séries produzidas são comparadas às séries esperadas, quantificando, assim, o desempenho e a acurácia do modelo. Neste trabalho, utilizam-se tais medidas para avaliar a capacidade de decomposição das abordagens discutidas nas seções anteriores, i.e., compara-se o componente determinístico original aos componentes determinísticos resultantes. Essa abordagem possibilita calcular o erro de decomposição. É importante notar que não há consenso entre os pesquisadores dessa área de conhecimento sobre qual o melhor critério de avaliação para cada cenário. Com o objetivo de abordar tal questão, foi realizado um estudo sobre as medidas comumente usadas na literatura (DUARTE *et al.*, 2019).

5.1 Distância de Minkowski, MAE, MSE e RMSE

A maneira mais simples de se comparar dois valores é calculando a diferença entre eles, dando assim uma ideia de distância ou erro (a depender do contexto) entre ambos os valores. Essa abordagem pode ser generalizada para o contexto de séries temporais, no qual a i -ésima observação da primeira série temporal x_i^1 é subtraída da i -ésima observação, correspondente no tempo, da segunda série temporal x_i^2 . Ao fim, calcula-se a soma de todos os erros individuais compondo a diferença entre duas séries temporais de mesmo tamanho.

No entanto, os erros positivos podem se anular com os negativos, levando a um resultado equivocado da diferença de uma série temporal para a outra. Esse problema pode ser verificado na Figura 26 em que uma série senoidal (pontos vazios) é comparada a uma série contante (pontos preenchidos). Observe que nesse cenário específico a diferença entre ambas as séries é, erroneamente, zero.

Figura 26 – As linhas vermelhas representam a diferença obtida a partir da subtração da série senoidal (pontos vazios) daquela de valores constantes (pontos pretos). A soma dessas diferenças (positivas e negativas) é igual a zero, levando a uma conclusão errônea.

Uma maneira simples de contornar este problema se dá pela soma dos módulos, ou o quadrado dos erros individuais calculados, evitando assim o efeito de cancelamento causado por sinais opostos. Essas abordagens levam a duas medidas populares, respectivamente: (i) distância de Manhattan e (ii) distância Euclidiana. A distância de Minkowski, por sua vez, é a generalização de ambas. Essa propicia diferentes interpretações a depender do valor de p ($p = 1$ para distância Manhattan e $p = 2$ para distância Euclidiana) escolhido de acordo com a Equação 5.1, em que X^1 e X^2 são as séries temporais que devem ser comparadas, e N é o número total de observações em cada série.

$$D(X^1, X^2) = \left(\sum_{i=1}^N |x_i^1 - x_i^2|^p \right)^{1/p} \quad (5.1)$$

Apesar de ambas as medidas utilizarem somente a magnitude e resolverem o problema de anulamento dos erros, a depender do número de observações da série, ao somar todos os erros, o valor final pode ser muito grande. Isso implica que séries temporais maiores tenderão a apresentar distâncias maiores entre si. Uma forma de resolver problema se dá pela divisão do erro pelo número de observações presentes na série. Neste contexto, surgem as medidas de erro: (i) Média do Erro Absoluto (do inglês *Mean Absolute Error* (MAE)) e (ii) Média do Quadrado do Erro (do inglês *Mean Squared Error* (MSE)). Um terceiro critério mais intuitivo e comum é a Raiz do Quadrado do Erro Médio (do inglês *Root Mean Squared Error* (RMSE)). Como a média de valores elevados ao quadrado pode assumir grande magnitude e, conseqüentemente, ser de difícil interpretabilidade, tipicamente emprega-se a raiz quadrada para normalizar o resultado final, diminuindo sua magnitude. Considere duas séries temporais X^1 e X^2 , ambas com N observações. As medidas MAE, MSE e RMSE entre ambas as séries podem ser calculadas pelas respectivas Equações 5.2, 5.3 e 5.4.

$$\text{MAE}(X^1, X^2) = \frac{1}{N} \sum_{i=1}^N |x_i^1 - x_i^2| \quad (5.2)$$

$$\text{MSE}(X^1, X^2) = \frac{1}{N} \sum_{i=1}^N (x_i^1 - x_i^2)^2 \quad (5.3)$$

$$\text{RMSE}(X^1, X^2) = \sqrt{\text{MSE}(x^1, x^2)} \quad (5.4)$$

Note que todos os resultados estão contidos no intervalo $[0; \infty)$, em que zero significa que as séries temporais comparadas são exatamente as mesmas. Como todas as medidas diferenciam as séries temporais única e exclusivamente baseadas na diferença entre os valores, é esperado que todas essas medidas avaliem o mesmo viés. A única distinção esperada está relacionada à magnitude dos resultados para a distância Euclidiana e a medida MSE pois ao elevarem o erro ao quadrado se tornam mais sensíveis a *outliers*.

5.2 Correlação de Pearson

As séries temporais também podem ser comparadas levando em consideração uma distância baseada em correlação. Neste contexto, um critério muito comum e usualmente utilizado é a correlação de Pearson (BENESTY *et al.*, 2009) calculada por meio da Equação 5.5 em que: X^1 e X^2 são as séries em comparação; x_i^1 e x_i^2 suas respectivas observações; $\bar{x}^1$ e $\bar{x}^2$ são as médias de todos os valores pertencentes a cada uma das séries; e N é o número total de observações de cada série.

$$\text{corr}(X^1, X^2) = \frac{\sum_{i=1}^N (x_i^1 - \bar{x}^1)(x_i^2 - \bar{x}^2)}{\sqrt{\sum_{i=1}^N (x_i^1 - \bar{x}^1)^2} \sqrt{\sum_{i=1}^N (x_i^2 - \bar{x}^2)^2}} \quad (5.5)$$

Essa medida é capaz de capturar a correlação linear entre as duas séries, fornecendo, ao fim, um valor contido no intervalo $[-1, 1]$. Esse valor serve para medir se ambas as séries apresentam o mesmo comportamento de tendência e sazonalidade. Se a correlação final for -1 , as séries têm uma correlação linear negativa, o que na prática implica que o comportamento de uma série é exatamente o complementar da outra, i.e., enquanto uma série cresce ao longo do tempo a outra decresce e vice-versa. Caso a correlação final seja 1 , há uma correlação positiva em que ambas as séries têm a mesma tendência e sazonalidade. Por fim, quando o valor calculado é 0 , não há correlação linear entre elas e não é possível identificar o comportamento de uma série com base na outra.

Neste ponto, é útil lembrar que, para ser considerada uma métrica, uma medida de distância $D(\cdot, \cdot)$ deve satisfazer as seguintes propriedades: i) (simetria) $D(t^1, t^2) = D(t^2, t^1)$; ii) (ser positiva) $D(t^1, t^2) \geq 0$; iii) (identidade) $D(t^1, t^2) = 0 \iff t^1 = t^2$; e iv) (desigualdade triangular) $D(t^1, t^2) \leq D(t^1, t^3) + D(t^2, t^3)$. Neste contexto, a correlação de Pearson não pode ser considerada uma métrica justamente por não atender a última propriedade (desigualdade triangular).

5.3 Dynamic Time Warping

A principal desvantagem apresentada pelas medidas anteriores é a dependência do alinhamento perfeito entre as séries temporais, i.e., um pequeno deslocamento entre as suas observações afeta completamente o resultado final. Para resolver esse problema, a medida *Dynamic Time Warping* (DTW) procura o melhor alinhamento não-linear entre as séries antes de calcular a distância (TORMENE *et al.*, 2009). Deste modo é possível encontrar padrões, tendências e sazonalidades com diferentes ritmos entre as séries.

Para entender o seu funcionamento, considere as séries temporais X^1 e X^2 e suas observações, na ordem, x_i^1 e x_j^2 . Como as séries não precisam ter o mesmo tamanho, considere N e M como o tamanho das séries 1 e 2, respectivamente. Para encontrar o melhor alinhamento entre as séries, a DTW utiliza uma matriz de suporte C com N linhas e M colunas. Assim, cada elemento $C[i, j]$ representa uma distância acumulada entre a observação x_i^1 e x_j^2 .

Para calcular o valor final de DTW, todos os elementos da matriz são inicializados com infinito, $C[:, :] = \infty$, e os dois primeiros elementos das séries são alinhados tendo a distância entre eles igual a zero, i.e., $C[0, 0] = 0$. Cada um dos demais termos da matriz são calculados tal como mostrado na Equação 5.6 de modo que, ao final, a distância entre os últimos elementos de ambas as séries, representado na matriz por $C[N, M]$, corresponde ao valor final de DTW.

$$C[i, j] = \text{dist}(x_i^1, x_j^2) + \min(C[i-1, j], C[i, j-1], C[i-1, j-1]) \quad (5.6)$$

Se pertinente, é possível calcular o *warping path*, o qual representa o melhor alinhamento entre as séries e que minimiza a medida DTW (MULLER, 2007). Para isso, calcula-se o caminho inverso, começando do elemento $C[M, N]$ e evoluindo em direção das células vizinhas que apresentam menor valor. O caminho é percorrido na matriz até alcançar o elemento $C[0, 0]$, indicando quais observações de cada série deve ser consideradas em diferentes instantes de tempo para obter o melhor encaixe entre elas. Ao fim, o *warping path* representa o melhor alinhamento entre as séries temporais X^1 e X^2 . Essa medida, se implementada em sua forma mais eficiente, tem complexidade computacional $O(MN)$ e, portanto, pode ser inviável para comparar duas séries com muitas observações.

5.4 Distância Média da Linha Diagonal

A maioria das medidas comumente adotadas para calcular a distância entre séries temporais avalia apenas as diferenças individuais entre pares de observações, sem considerar, contudo, o comportamento global que pode apresentar tendências e/ou sazonalidades. Ainda, se houver uma diferença de amostragem e as séries não apresentarem o mesmo tamanho ou se houver um deslocamento temporal, as medidas utilizadas apresentaram comportamento e valores incoerentes.

Para resolver esse problema, [Rios e Mello \(2013\)](#) projetaram uma nova medida chamada Distância Média da Linha Diagonal (do inglês *Mean Distance from the Diagonal Line* (MDDL)). Baseada na DTW ([TORMENE et al., 2009](#))¹, a MDDL inicia calculando o *warping path* entre as duas séries temporais em avaliação. Seu resultado é obtido calculando a área absoluta entre o *warping path* e a linha diagonal da matriz C (vide Seção 5.3). O valor final da medida é definido no intervalo $\left[0, \int_a^b |f(x)|dx\right]$, tal que $f(x)$ é a função de MDDL usada para calcular o *warping path*, a e b são o primeiro e último elementos. Ao considerar esse intervalo, é possível notar que os valores de MDDL variam de zero, representando uma correspondência perfeita entre as séries, até a área máxima formada pelo *warping path* e a linha diagonal. Na pior das hipóteses, quando as observações previstas e esperadas são completamente diferentes, MDDL retorna a área do triângulo isósceles, cujos lados são formados por N esperado e N de observações previstas, enquanto a hipotenusa é formada pela linha diagonal. Consequentemente, esse intervalo pode ser escrito como $\left[0, +\frac{(N^2-N)}{2}\right]$. Em suma, o número de observações é finito, assim a distância Euclidiana entre o *warping path* e a linha diagonal pode ser usada para calcular a valor de MDDL ([RIOS; MELLO, 2013](#)).

A fim de ilustrar o funcionamento da medida MDDL, considere duas séries temporais: (i) uma senoidal e (ii) uma série formada puramente por um ruído branco, seguindo uma distribuição de probabilidades Normal com média 0 e variância 0,5². Ao se calcular a MDDL entre a série puramente senoidal e o ruído branco (Figura 27(a)) obtêm-se o resultado 4,11. De fato, o baixo valor de MDDL é esperado, pois a adição do ruído modificou a magnitude das observações, mas manteve o comportamento global com relação a tendências e sazonalidades. Esse fato pode ser observado pela pequena área formada entre o *warping path* e a linha diagonal (Figura 27(c)). Apenas como referência, ao se comparar a série senoidal com o ruído branco (Figura 27(b)), obtêm-se um valor para MDDL igual a 69,78. Essa alta magnitude pode ser verificada na Figura 27(d) que mostra a grande área formada entre o *warping path* e a linha diagonal.

5.5 Distância média dos Atratores

Todas as medidas até aqui mencionadas trabalham única e exclusivamente sobre o espaço tempo. No entanto, como o objetivo é decompor a série temporal preservando, contudo, o viés do componente determinístico no espaço-fase, é necessário um critério que seja capaz de avaliar dois atratores: o predito e o observado. Neste contexto, foi proposta neste doutorado uma medida denominada Distância Média dos Atratores (do inglês *Mean Distance from Attractors* (MDA)), a qual considera a imersão de duas séries temporais, X^1 e X^2 , no espaço-fase utilizando o teorema de [Takens \(1981\)](#). Para isso, o usuário precisa informar a dimensão de embutir m e a de separação τ , a fim de desdobrar as relações entre as observações da série no espaço-fase (uma possível metodologia para estimar tais parâmetros é descrita nas Seções 2.3.3 e 2.3.4).

¹ Devido à essa dependência, a MDDL também não pode ser considerado uma métrica.

Figura 27 – Em (a), tem-se uma série senoidal preta e a mesma adicionada de ruído em vermelho. A área entre o *warping path* e a linha diagonal é quase desprezível (b). Ao contrário, ao se comparar uma série senoidal (preta) com uma série aleatória (vermelha) presente em (c), tem-se uma área significativa tal como visto em (d).

Após a imersão, calcula-se a distância Euclidiana $d(.,.)$ entre cada par de estados sob o mesmo índice dos atratores em comparação, i.e., a distância entre o primeiro estado do atrator S^1 e o primeiro do atrator S^2 , e assim por diante. Por fim, calcula-se a média de todas essas distâncias que então é atribuída como a medida MDA entre o par de atratores. Como consequência da metodologia aplicada, quanto menor o valor de MDA mais similares são ambos atratores. Além disso, outra característica pertinente é que tal medida é capaz de comparar recorrências, tendências e sazonalidades no espaço-fase. Nesse sentido, se dois atratores apresentam o mesmo formato, então necessariamente estes terão o mesmo comportamento no espaço tempo. Entretanto, para que tudo funcione, ambas as séries temporais em comparação devem necessariamente ter o mesmo número de observações, o que pode ser um problema para técnicas de decomposição que produzem componentes com tamanhos distintos da série original.

As diversas medidas até agora apresentadas são comumente encontradas nos diversos trabalhos da área de decomposição de séries temporais. Há uma grande dificuldade dos pesquisadores em determinar qual o melhor critério para avaliar suas respectivas técnicas. Por

esse motivo, a grande maioria dos trabalhos utiliza mais de uma medida com o objetivo de tirar qualquer viés de seleção ou de avaliação. No entanto, cada medida mensura um viés específico da decomposição e a utilização de várias que avaliam o mesmo aspecto não contribui para diminuir o viés de seleção. Neste contexto, para determinar a melhor forma de mensurar os resultados da decomposição, foram inicialmente executados experimentos em bases de dados sintéticas.

5.6 Bases de Dados Sintéticas

Existem várias estratégias para avaliar o processo de decomposição, as quais podem ser selecionadas de acordo com conhecimentos prévios ou conclusões extraídas do próprio comportamento dos dados ou domínio de aplicação (KANTZ; SCHREIBER, 2004). No entanto, a seleção de uma única estratégia é muito complexa ou, eventualmente, inviável para avaliar as decomposições de propósito geral. Com o objetivo de melhor entender essas estratégias, foram realizados vários experimentos para apoiar o processo decisório da abordagem mais adequada para certos cenários de interesse. Para isso, 55 séries temporais sintéticas foram produzidas a partir da combinação de cinco componentes determinísticos e onze componentes estocásticos. Cada um com uma forma específica de atrator no espaço-fase, os componentes determinísticos escolhidos descrevem fenômenos do mundo real e têm sido amplamente utilizados na literatura (ALLIGOOD; SAUER; YORKE, 1996; KANTZ; SCHREIBER, 2004; PAGLIOSA; MELLO, 2017).

Muito usado em modelos climáticos, como temperaturas (ZHANG *et al.*, 2017), as funções senoidais são caracterizadas por atratores elípticos. Esse comportamento cíclico no formato do atrator é causado pela sazonalidade da série no espaço tempo. Quanto maior o ciclo, maior será o período da sazonalidade da série. O sistema de Lorenz, e sua simplificação através do sistema de Rössler, descreve a dinâmica de um fluido sobre uma superfície uniformemente aquecida na base e resfriada uniformemente no topo. Essa dinâmica confere a ambos os atratores transições suaves entre os estados e diversos traçados paralelos. Todo esse processo de aquecimento e resfriamento confere ao sistema uma periodicidade visível ao formato do atrator em que Lorenz se assemelha ao símbolo do infinito e Rössler a um disco com um dos lados levantado na terceira dimensão. Finalmente, os mapas de Hénon e Logístico foram amplamente estudados devido às suas baixas dimensões e características de caoticidade. Ambos têm atratores com transições abruptas e formas que lembram uma parábola côncava. A Figura 28 mostra tais componentes determinísticos (coluna da esquerda) e seus respectivos atratores no espaços de fase (coluna da direita).

Cada um dos componentes determinísticos, com mil observações, foi normalizado no intervalo $[0; 1]$ e somado a outros 11 componentes estocásticos, também com mil observações cada, formando ao todo 55 cenários experimentais. O grupo de séries estocásticas é formado por: uma série nula (todos seus valores são iguais a zero), cinco séries formadas por um distribuição

Figura 28 – Séries temporais determinísticas e seus respectivos atratores no espaço-fase. Em sequência: Seno, Lorenz, Rössler, Mapa Logístico e Hénon.

Fonte: Elaborada pelo autor.

de probabilidades Normal com média zero e variância $\{0,01^2; 0,05^2; 0,10^2; 0,15^2; 0,20^2\}$, e outras cinco séries formadas por observações que seguem uma distribuição de probabilidades uniforme com valor mínimos e máximos iguais a: $\{-0,01; 0,01\}$, $\{-0,05; 0,05\}$, $\{-0,10; 0,10\}$, $\{-0,15; 0,15\}$ e $\{-0,20; 0,20\}$.

A combinação dos componentes determinísticos com os estocásticos forma uma bateria de testes com 55 séries temporais distintas. Com o objetivo de tornar os resultados mais robustos e confiáveis, 30 baterias de testes foram geradas e utilizadas no processo de decomposição. Em comum, todas as baterias utilizam os mesmos componentes determinísticos, porém somados aos componentes estocásticos e aleatórios distintos. A utilização das bases sintéticas é fundamental para determinar a capacidade de decomposição das estratégias, pois permitem comparação com componentes conhecidos. Assim, é preciso determinar qual a melhor medida para distinguir e melhor avaliar as técnicas de decomposição.

5.7 Comparação

Para determinar qual a melhor forma de mensurar os resultados da decomposição, as 30 baterias de testes (cada uma com 55 séries temporais) foram submetidas às técnicas do estado da arte (Capítulo 3) variando os seus hiper-parâmetros. Essa combinação resultou em 2,3 milhões de experimentos que foram medidos utilizando as metodologias descritas neste capítulo. Os resultados de cada medida foram comparados utilizando a correlação de Pearson para verificar quais delas estavam avaliando o mesmo viés de decomposição. O resultado final é sumarizado na Figura 29.

	MAE	RMSE	Minkowski P1	Minkowski P2	Minkowski P3	Pearson Corr.	MDDL	MDA
MAE	1.0000	0.9981	1.0000	0.9981	0.9944	-0.9038	0.3858	0.9930
RMSE	0.9981	1.0000	0.9981	1.0000	0.9990	-0.9057	0.3921	0.9951
Minkowski P1	1.0000	0.9981	1.0000	0.9981	0.9944	-0.9038	0.3858	0.9930
Minkowski P2	0.9981	1.0000	0.9981	1.0000	0.9990	-0.9057	0.3921	0.9951
Minkowski P3	0.9944	0.9990	0.9944	0.9990	1.0000	-0.9036	0.3944	0.9943
Pearson Corr.	-0.9038	-0.9057	-0.9038	-0.9057	-0.9036	1.0000	-0.5478	-0.9216
MDDL	0.3858	0.3921	0.3858	0.3921	0.3944	-0.5478	1.0000	0.4303
MDA	0.9930	0.9951	0.9930	0.9951	0.9943	-0.9216	0.4303	1.0000

Figura 29 – Matriz de Correlação calculada para pares de medidas.

Como as medidas MAE, RMSE, Minkowski P1, Minkowski P2 e Minkowski P3 avaliam as séries levando em consideração somente a diferença entre suas observações, todas acabam tendo o mesmo viés de avaliação. Esse cenário é ainda mais evidente em nossos testes pois todas as séries foram normalizadas no intervalo $[0; 1]$ e a etapa de decomposição forneceu valores muito próximos aos esperados. Assim, apesar das medidas RMSE, Minkowski P2 e Minkowski P3 serem sensíveis a *outliers*, essa discrepância no valor final da medida não ficou evidente.

As medidas MDA e correlação de Pearson, que são tipicamente utilizadas para capturar se as tendências e sazonalidades estão sendo obedecidas, também apresentaram alta correlação com as medidas citadas anteriormente. Isso demonstra que as simples diferenças entre as observações já são suficientes para mensurar todo o comportamento de uma série temporal com relação à outra. Cabe aqui ressaltar que como a correlação de Pearson é uma medida de similaridade, sua correlação com as demais medidas ficou negativa mas, ainda sim, com alta magnitude. Diferentemente, a medida MDDL apresentou uma baixa correlação com todas as demais medidas, o que se deve, principalmente, ao fato dela tentar primeiramente encontrar o melhor encaixe temporal das séries antes de proceder com as comparações. De acordo com nossas análises realizadas a partir da Figura 29, as medidas MDDL e MAE são suficientes para avaliar o processo de decomposição e analisar diferentes vieses.

5.8 Considerações Finais

Há na literatura diversas medidas de avaliação para os algoritmos de decomposição. No entanto, a falta de um consenso entre os pesquisadores sobre qual a melhor forma de mensurar as metodologias faz com que a maioria empregue a maior quantidade possível de medidas sem avaliar se elas sumarizam o mesmo viés. Como consequência, pode-se chegar a conclusões distorcidas pela análise de diferentes critérios que, na verdade, capturam o mesmo viés. Neste contexto, este capítulo apresentou um estudo sobre as medidas mais utilizadas e viáveis para avaliar decomposições de séries temporais. Os resultados mostram que as medidas MDDL e MAE são suficientes para medir vieses distintos das decomposições consideradas.

COMPARAÇÃO ENTRE TÉCNICAS DE DECOMPOSIÇÃO DE SÉRIES TEMPORAIS

O principal objetivo deste trabalho foi desenvolver uma abordagem para decompor séries temporais em seus respectivos componentes determinístico e estocástico, preservando, contudo, o viés determinístico original da série. Esperou-se, assim, obter melhores resultados de modelagem e, conseqüentemente, de predição de observações. A validação deste objetivo foi dividida em duas etapas: (i) validação das novas abordagens de decomposição e (ii) validação da hipótese de melhoria na etapa de predição da série. Para a primeira, foco deste capítulo, as principais técnicas de decomposição do estado-da-arte (vide Capítulo 3) bem como as três aqui propostas (vide Capítulo 4) foram submetidas a diversas avaliações visuais e quantitativas. Desse modo, analisou-se o resultado das técnicas de decomposição, bem como quais seus melhores cenários de aplicação, guiando o usuário no processo de decomposição.

6.1 Avaliação Visual

Para compreender o comportamento das técnicas de decomposição com relação aos diversos cenários de teste, foram inicialmente realizadas avaliações visuais. Cada um dos cinco processos determinísticos geradores (Seno, Lorenz, Rössler, Mapa Logístico e Hénon) foi somado a um ruído, que segue uma distribuição de probabilidades Normal, e submetidos às 11 diferentes técnicas de decomposição: 1) Transformada de Fourier – FT (Seção 3.1), 2) Transformada Wavelet – WT (Seção 3.2), 3) Média Móvel – MA (Seção 3.3), 4) Análise Espectral Singular – SSA (Seção 3.4), 5) Lazy (Seção 3.6), 6) GHKSS (Seção 3.6), 7) *Empirical Mode Decomposition* utilizando *Recurrence Plot* – EMD-RP (Seção 3.5.1), 8) *Empirical Mode Decomposition* utilizando *Mutual Information* – EMD-MI (Seção 3.5.2), 9) **Spring** (Seção 4.1), 10) **Spring Time Domain** – **Spring-TD** (Seção 4.2) e 11) **Spring*** (Seção 4.3).

As séries determinísticas resultantes desse processo foram imersas em seus respecti-

vos espaços-fase, cujos atratores são ilustrados na Figura 30. Cada coluna representa um dos cinco processos determinísticos geradores e cada linha representa uma das onze técnicas de decomposição. Com intuito de facilitar a visualização, comparação e interpretação dos resultados, particularmente a primeira linha (verde) contém os atratores sem a adição do componente estocástico e, portanto, o formato esperado. As oito linhas seguintes (laranja) contêm as técnicas de decomposição do estado da arte, e as três últimas (azul), as técnicas **Spring**, **Spring-TD** e **Spring***, respectivamente. Todos os espaços-fase da Figura 30 estão no intervalo $[-1, 1]$ para que seja possível observar o comportamento de técnicas que diminuem a amplitude da série temporal no processo de decomposição e, consequentemente, produzem atratores mais pontuais, como é o caso da EMD-MI. Assim, o processo de avaliação visual consiste em comparar o formato do atrator resultante com o formato do atrator na primeira linha. Ainda, é possível comparar a qualidade da decomposição entre as técnicas e observar possíveis vieses.

Por inspeção visual, nota-se um bom funcionamento de todas as técnicas de decomposição para a série com componente determinístico senoidal. De fato, esse comportamento era esperado uma vez que a maioria das técnicas do estado da arte impõem um viés senoidal no componente determinístico resultante, o que coincide com o viés original da série. Bons resultados são também observados para o atrator de Lorenz, com uma ressalva para o comportamento da técnica EMD-MI. De modo geral, todas as técnicas conseguiram manter a topologia do atrator, preservando o buraco característico da alça superior. Em destaque, tem-se a técnica SSA que removeu a cintura que separa ambas as alças do atrator e as técnicas Lazy e **Spring** que afinaram a alça superior.

No que se diz respeito ao atrator de Rössler, o resultado foi bem distinto entre as técnicas. As transformadas de Fourier e Wavelet distorceram completamente os atratores o que, consequentemente, produziu um componente determinístico totalmente diferente do original. As técnicas EMD-MI e EMD-RP produzem uma decomposição com uma grande perda de amplitude do componente determinístico e, por este motivo, têm o atrator reduzido e denso. As técnicas MA, SSA, **Spring** e **Spring-TD** destroem a terceira dimensão característica do atrator, mantendo, contudo, a topologia e a característica do atrator. Por fim, os melhores resultados visuais foram obtidos com as técnicas Lazy, GHKSS e **Spring***.

Para o Mapa Logístico e Hénon os resultados são bem similares, ou seja, as técnicas WT, MA, SSA, EMD-MI e **Spring-TD** distorcem completamente os atratores. Isso ocorre pois séries de natureza caótica têm altas frequências e comportamento que se assemelha ao estocástico, os quais são filtrados por essas técnicas. Como Lazy, GHKSS, **Spring** e **Spring*** trabalham sobre o espaço-fase, elas conseguem melhor preservar o formato do atrator e, por consequência, o componente determinístico. FT não conseguiu decompor o atrator e o resultado final é idêntico ao atrator inicial. Por fim, o péssimo resultado obtido pela técnica **Spring-TD** para os dois últimos cenários se dá devido à forma como ela opera sobre o espaço-fase e como esses atratores são construídos. As transições entre estados consecutivos, i.e., s_i e s_{i+1} , não são suaves nem

		Mapa				
		Seno	Lorenz	Rössler	Logístico	Hénon
Ground Truth
	FT
	WT
	MA
	SSA
	Lazy
	GHKSS
	EMD-MI
	EMD-RP
	Spring
	Spring-TD
	Spring*

Figura 30 – Matriz para avaliação Visual: cada coluna representa um componente determinístico e cada linha uma técnica de decomposição. A comparação deve ser feita observando a proximidade de cada atrator com seu atrator referência na primeira linha. Cabe destacar os bons resultados apresentado pelas técnicas **Spring**, **Spring-NP**, **Lazy** e **GHKSS**

Fonte: Elaborada pelo autor.

próximas e, por esse motivo, a técnica **Spring-TD** deforma o espaço-fase. Esse fato corrobora com a afirmação de que a **Spring-TD** deve ser aplicada somente sobre atratores com transições suaves entre seus estados (vide Seção 4.2).

6.2 Avaliação Quantitativa

A etapa visual, apesar de propiciar uma boa avaliação, permite apenas conclusões subjetivas, sendo difícil distinguir quando ocorrem pequenas diferenças entre os resultados. Para confirmar as conclusões visuais e remover qualquer abstração, há a necessidade de uma avaliação quantitativa, a qual considerou 30 baterias de testes, cada uma contendo as 55 séries temporais, tal como descrito na Seção 5.6. Essas séries foram fornecidas aos 11 algoritmos de decomposição e, após o processamento, o componente determinístico encontrado foi comparado ao esperado, utilizando as medidas MAE e MDDL — conforme definido no Capítulo 5. As médias dos resultados são apresentadas nas Tabelas 2 e 3. As oito primeiras linhas mostram os resultados obtidos pelas técnicas que compõem o estado da arte (FT, WT, MA, SSA, Lazy, GHKSS, EMD-MI e EMD-RP) e as três últimas são associadas às técnicas **Spring**, **Spring-TD** e **Spring***. Por sua vez, cada coluna representa um dos cinco cenários de teste (Seno, Lorenz, Rössler, Mapa Logístico e Hénon).

Os resultados da medida MAE (Tabela 2) corroboram com aqueles obtidos via inspeção visual, descrita na seção anterior, e demonstram superioridade das técnicas **Spring** e **Spring*** em termos das informações consolidadas. Para a série temporal com viés senoidal, destaque para as técnicas pertencentes ao atual estado da arte com bons resultados. Já para os cenários não lineares, destacam-se **Spring** e suas variações. Os resultados são também similares para o cenário de Lorenz, em que a técnica GHKSS tem o menor valor de MAE. As maiores diferenças entre resultados acontecem para os cenários com dados oriundos de Rössler, Mapa Logístico e Hénon. Nesses casos, as técnicas mais tradicionais — FT, WT, MA e SSA — apresentam altos valores de MAE o que faz com que, no consolidado, elas apresentem resultados significativamente inferiores às demais. Por fim, as técnicas Lazy, GHKSS, **Spring** e **Spring*** apresentam os melhores resultados.

Os resultados da medida MDDL são complementares aos da MAE, pois aquela é eficiente para mensurar semelhanças entre duas séries dada a melhor sincronização entre ambas. Interessante notar que, apesar de seguir uma tendência em que as melhores técnicas com relação à MAE são também as melhores em termos de MDDL, ao fim, na média, as melhores técnicas com relação a MDDL são WT, MA e **Spring-TD**. Em particular, no contexto da série senoidal, as técnicas FT, WT, GHKSS e **Spring** apresentam os melhores resultados, com destaque para FT e **Spring**. Em Lorenz, as técnicas FT e GHKSS perdem desempenho deixando WT e **Spring** como as mais indicadas. No cenário de Rössler, como verificado na avaliação visual (Seção 6.1), algumas técnicas distorceram a terceira dimensão característica do atrator, fato que pode ser

Tabela 2 – Média e desvio padrão da medida MAE para cada técnica em cada cenário determinístico de teste. Destaque para as técnicas **Spring** e **Spring***, que apresentaram os melhores resultados em termos das informações consolidadas.

	Seno	Lorenz	Rössler	Mapa Logístico	Hénon	Total
FT	0,04 ± 0,04	0,07 ± 0,04	0,22 ± 0,07	0,08 ± 0,07	0,07 ± 0,06	0,09 ± 0,08
WT	0,05 ± 0,04	0,10 ± 0,05	0,22 ± 0,05	0,53 ± 0,00	0,39 ± 0,01	0,26 ± 0,18
MA	0,03 ± 0,02	0,09 ± 0,05	0,20 ± 0,07	0,50 ± 0,02	0,40 ± 0,02	0,24 ± 0,18
SSA	0,02 ± 0,02	0,07 ± 0,03	0,13 ± 0,02	0,24 ± 0,04	0,24 ± 0,04	0,14 ± 0,09
Lazy	0,05 ± 0,04	0,08 ± 0,03	0,09 ± 0,04	0,06 ± 0,04	0,06 ± 0,04	0,07 ± 0,04
GHKSS	0,05 ± 0,05	0,05 ± 0,04	0,06 ± 0,04	0,06 ± 0,05	0,06 ± 0,05	0,06 ± 0,05
EMD-MI	0,13 ± 0,22	0,26 ± 0,02	0,39 ± 0,00	0,55 ± 0,00	0,48 ± 0,00	0,36 ± 0,18
EMD-RP	0,03 ± 0,03	0,08 ± 0,06	0,30 ± 0,13	0,07 ± 0,08	0,06 ± 0,06	0,11 ± 0,13
Spring	0,04 ± 0,04	0,06 ± 0,03	0,09 ± 0,03	0,04 ± 0,03	0,05 ± 0,03	0,06 ± 0,04
Spring-TD	0,03 ± 0,03	0,07 ± 0,03	0,19 ± 0,05	0,49 ± 0,01	0,41 ± 0,01	0,24 ± 0,19
Spring*	0,05 ± 0,05	0,06 ± 0,05	0,06 ± 0,05	0,06 ± 0,03	0,06 ± 0,04	0,06 ± 0,04

comprovado nos altos valores de MDDL. Ainda, altos valores de MDDL corroboram com a inspeção visual para técnicas que produzem componentes determinísticos com perda de amplitude. Para as séries oriundas do Mapa Logístico e de Hénon, a **Spring** obteve os melhores resultados.

Tabela 3 – Média e desvio padrão da medida MDDL para cada técnica em cada cenário determinístico de teste. Destaque para as técnicas **Spring-TD**, **WT** e **MA** que apresentaram os melhores resultados em termos das informações consolidadas.

	Seno	Lorenz	Rössler	Mapa Logístico	Hénon	Total
FT	1,54 ± 0,65	3,08 ± 1,20	12,58 ± 4,03	1,92 ± 0,83	1,18 ± 0,65	4,06 ± 4,73
WT	1,66 ± 0,70	2,57 ± 0,72	5,74 ± 2,60	3,66 ± 1,13	3,65 ± 0,50	3,46 ± 1,93
MA	2,33 ± 0,75	3,63 ± 0,77	4,64 ± 2,40	4,21 ± 1,62	3,23 ± 0,68	3,61 ± 1,62
SSA	2,31 ± 1,75	3,44 ± 0,91	22,12 ± 2,71	5,40 ± 7,26	3,83 ± 0,66	7,42 ± 8,24
Lazy	2,15 ± 0,54	3,26 ± 1,20	23,29 ± 2,40	2,14 ± 0,92	1,30 ± 0,90	6,43 ± 8,56
GHKSS	1,63 ± 0,82	3,94 ± 1,74	24,88 ± 9,57	1,90 ± 1,16	1,30 ± 1,02	6,73 ± 10,13
EMD-MI	37,99 ± 72,38	11,95 ± 3,55	67,80 ± 31,14	39,04 ± 21,00	53,38 ± 28,63	42,03 ± 42,87
EMD-RP	2,87 ± 1,34	5,21 ± 2,38	18,97 ± 7,95	2,82 ± 1,43	2,05 ± 1,55	6,38 ± 7,46
Spring	1,54 ± 0,70	2,95 ± 1,07	17,34 ± 3,74	1,58 ± 0,92	1,14 ± 0,63	4,91 ± 6,51
Spring-TD	2,18 ± 1,08	3,10 ± 0,97	4,53 ± 2,62	3,50 ± 1,19	3,20 ± 0,51	3,30 ± 1,64
Spring*	2,59 ± 1,40	6,51 ± 2,88	31,02 ± 11,25	3,26 ± 0,90	2,25 ± 1,36	9,13 ± 12,25

De modo geral, chama a atenção os resultados obtido pela EMD-MI, a qual é, até então, a única das abordagens de decomposição de séries temporais que não necessita de parametrização explícita do usuário e, consequentemente, conhecimento do domínio e da técnica. Seus resultados sub-ótimos são, contudo, o preço a ser pago para tal habilidade. Neste contexto, a técnica **Spring*** se coloca no estado da arte como a melhor abordagem de decomposição sem parâmetros. Seus resultados são equiparáveis às técnicas parametrizáveis e, em alguns cenários, até mesmo superiores.

Apesar de diferentes técnicas se destacarem para os diferentes cenários de teste, o uso somente de médias para avaliação pode levar a conclusões equivocadas. Isso porque, caso alguma técnica apresente grandes valores (alta amplitude) em um determinado cenário de teste, o valor final da média será alterado, levando à conclusão que tal técnica não apresenta bons resultados

de maneira global. Faz-se necessária, portanto, uma análise complementar. Neste contexto, os resultados de todas as baterias de teste foram coletados e são representados nos gráficos de violino (do inglês *Violin plot*) na Figura 31.

Figura 31 – Gráfico de Violino com resultados da medida MAE (a) e MDDL (b). É possível observar uma dispersão nos resultados sendo que a melhor técnica é aquela que produz a maior concentração em valores menores.

Fonte: Elaborada pelo autor.

Esses gráficos mostram a distribuição dos resultados para cada técnica em todos os testes. Nesse sentido, a melhor técnica é aquela que apresenta uma maior concentração em menores valores e, consequentemente, um violino com alta concentração na base. Em particular, é possível observar que os resultados seguem uma distribuição multimodal, consequência de diferentes resultados para cada cenário determinístico submetido à avaliação. Essa característica corrobora o fato de que a análise somente baseada em médias pode levar a conclusões equivocadas.

No que diz respeito à medida MAE (Figura 31a), as técnicas que apresentam menor dispersão são a Lazy, GHKSS, **Spring** e **Spring***, sugerindo uma consistência nos resultados, independente das características da série em estudo. O destaque fica para as técnicas GHKSS e **Spring** que apresentam maior concentração em valores menores e bases mais largas. No contexto da MDDL (Figura 31b), o destaque fica para a **Spring**, a qual apresenta a maior concentração em valores menores, e para WT, MA e **Spring-TD** que têm uma menor dispersão em MDDL. Por fim, analisando ambos gráficos, a técnica **Spring** é tida como a melhor em termos das informações consolidadas. Ao se comprar as abordagens não paramétricas, mais uma vez, a **Spring-TD** supera os resultados da EMD-MI que conta com alta dispersão nos resultados.

Embora os gráficos de violino complementem a análise, eles ainda são bastante subjetivos, dando apenas uma visualização da distribuição dos resultados. Note que não é possível ver em qual cenário uma técnica específica é melhor que outra. Para isso, a Tabela da Figura 32 endereça tal necessidade, mostrando o percentual das vezes em que cada técnica (linha) supera o resultado das demais (colunas). De modo geral, as técnicas **Spring** e GHKSS obtiveram os melhores resultados tanto para a medida MAE quanto para MDDL. Em particular, destaque para **Spring**, que obteve melhores resultados na maioria das vezes com relação as medidas MAE e MDDL. Ainda, a técnica **Spring*** superou EMD-MI com relação ao MAE (86%) e também quanto à MDDL (93%), além de produzir bons resultados frente às demais técnicas em termos de MAE. Isso a coloca como a melhor abordagem não paramétrica para decomposição de séries temporais.

A fim de fornecer garantias sobre a validade e não-aleatoriedade dos resultados, seguiu-se a abordagem proposta por Demsar (2006), que aplica o teste de hipóteses não paramétrico Wilcoxon para validar a diferença nos resultados. Segundo esse teste, a hipótese nula afirma que a diferença entre os pares de resultados segue uma distribuição simétrica em torno de zero. Isso significa que, ao obter um p -value pequeno e, consequentemente, rejeitar a hipótese nula, a diferença entre as variáveis de interesse é significativa. Considerando as comparações por pares, obtêm-se valores de p -value menores que 10^{-3} , confirmando que **Spring** obteve os melhores resultados.

6.3 Considerações Finais

Diversos algoritmos e metodologias foram desenvolvidos com o propósito de decompor séries temporais em seus respectivos componentes estocásticos e determinísticos. Essas se diferenciam em termos do viés imposto sobre a decomposição, a metodologia da filtragem, e o espaço sobre o qual trabalham. Essas diferenças proporcionam, ao fim, resultados distintos e, consequentemente, diferenças nas medidas de avaliação. Neste sentido, este capítulo reportou uma avaliação sobre as principais técnicas do estado da arte tanto em termos da interpretação visual quanto quantitativa. Essas avaliações permitiram observar o comportamento e o viés de cada abordagem e concluir sobre seus melhores usos. Resultados comprovam a superioridade da

	FT	WT	MA	SSA	Lazy	GHKSS	EMD-MI	EMD-RP	Spring	Spring-TD	Spring*
FT	0%	82%	70%	55%	44%	31%	89%	38%	30%	66%	36%
WT	18%	0%	40%	9%	18%	14%	88%	26%	11%	36%	14%
MA	30%	60%	0%	14%	24%	16%	86%	33%	17%	42%	19%
SSA	45%	91%	86%	0%	36%	26%	96%	43%	26%	81%	27%
Lazy	56%	82%	76%	64%	0%	30%	85%	50%	21%	67%	38%
GHKSS	69%	86%	84%	74%	70%	0%	86%	71%	49%	78%	58%
EMD-MI	11%	12%	14%	4%	15%	14%	0%	24%	14%	12%	14%
EMD-RP	65%	74%	67%	57%	50%	29%	80%	0%	35%	64%	38%
Spring	70%	89%	83%	74%	79%	51%	86%	65%	0%	78%	63%
Spring-TD	34%	64%	58%	19%	33%	22%	88%	36%	22%	0%	24%
Spring*	64%	86%	81%	73%	62%	42%	86%	63%	37%	76%	0%

(a) Comparação entre as técnicas com relação a métrica MAE.

	FT	WT	MA	SSA	Lazy	GHKSS	EMD-MI	EMD-RP	Spring	Spring-TD	Spring*
FT	0%	59%	74%	84%	71%	60%	100%	84%	46%	66%	90%
WT	39%	0%	61%	75%	52%	44%	100%	62%	38%	45%	61%
MA	26%	39%	0%	63%	37%	33%	99%	55%	25%	31%	54%
SSA	14%	23%	37%	0%	33%	35%	97%	46%	15%	28%	56%
Lazy	25%	48%	63%	67%	0%	43%	98%	69%	19%	57%	83%
GHKSS	29%	54%	67%	63%	53%	0%	99%	72%	30%	63%	88%
EMD-MI	0%	0%	1%	3%	2%	1%	0%	1%	0%	1%	7%
EMD-RP	2%	37%	45%	52%	26%	17%	95%	0%	15%	38%	63%
Spring	50%	60%	75%	83%	80%	65%	100%	81%	0%	69%	93%
Spring-TD	32%	53%	69%	71%	43%	36%	99%	60%	29%	0%	60%
Spring*	4%	37%	46%	42%	17%	7%	93%	30%	6%	39%	0%

(b) Comparação entre as técnicas com relação a métrica MDDL.

Figura 32 – Comparação par-a-par das técnicas de decomposição. Percentual das vezes em que uma dada técnica (linha) supera o resultado das demais (colunas) com relação às medidas MAE (a) e MDDL (b).

Fonte: Elaborada pelo autor.

Spring que, na maior parte dos cenários, apresentou resultados superiores às demais técnicas. Destaque, também, para a **Spring*** que se colocou com a melhor estratégia sem parâmetros ao superar EMD-MI na maior parte dos experimentos realizados.

PREDIÇÃO DE SÉRIES TEMPORAIS

Apesar da técnica **Spring** e suas variações apresentarem bons resultados para decompor séries temporais em seus respectivos componentes determinístico e estocástico, a hipótese levantada nesta tese de doutorado é que ao preservar o viés determinístico após a etapa de decomposição, haverá melhoras no processo de predição. Para finalmente validar esta hipótese, além de testar a limitação das técnicas de decomposição, três bases de dados reais foram utilizadas para proceder com predições: S&P 500 (Seção 7.1.1), *Sunspots* Mensal (Seção 7.1.2) e Circuito de Chua (Seção 7.1.3). Elas foram escolhidas por apresentarem atratores com características que permitem compreender as vantagens e limitações da **Spring** e suas variantes.

7.1 Séries Temporais Reais

7.1.1 S&P 500

Um dos mercados que mais utilizam a análise e predição de séries temporais é o financeiro. Aplicações como previsão de indicadores macro econômicos e previsão de câmbio estão entre as de maior interesse, pois apresentam potencial de ganhos significativos de capital para seus investidores. Neste contexto, encontra-se o índice S&P 500, o qual é formado pelas 500 principais ações¹ negociadas na bolsa de valores de Nova Iorque (do inglês New York Stock Exchange (NYSE)) e pela bolsa automática Nasdaq. Esse índice é tipicamente utilizado para compreender o comportamento do mercado de ações, da economia americana e mundial.

Ao analisar o valor do índice entre agosto de 2013 e dezembro de 2017 (Figura 33a), observa-se um comportamento que, aparentemente, não segue um padrão. De fato, as séries financeiras estão sujeitas a diversas influências externas, tais como oriundas do cenário político, clima, e resultados financeiros das empresas que compõem o índice, o que torna a tarefa de

¹ Importância da ação mensurada pelo tamanho de mercado, liquidez e representação de grupo industrial.

predição muito complexa. A consequência desta falta de padrão pode ser observada no formato do atrator desse série em seu respectivo espaço-fase (Figura 33b).

(a) Série histórica S&P 500.

(b) Atrator da série temporal S&P 500.

Figura 33 – Série histórica do valor de abertura da S&P 500 de agosto/2013 a dezembro/2017. Em (a) série temporal com os valores de abertura e em (b) atrator da série imerso no espaço-fase.

Fonte: Elaborada pelo autor.

De fato, a ausência de padrão e a influência de fatores externos tornam o comportamento da série praticamente aleatório. Neste cenário, é esperado que os algoritmos de decomposição tenham dificuldades para melhorar os resultados de predição, fato que motiva seu emprego para explorar limitações das abordagens propostas nesta tese.

7.1.2 *Sunspots Mensal*

Corpos celestes, tais como planetas e sóis, são influenciados por fenômenos sazonais, sejam devido à rotação, translação ou qualquer outra característica do corpo em si. Logo, o estudo da série temporal desses comportamentos permite auxiliar na compreensão da dinâmica do astro que, a depender, tem grandes implicações sobre a dinâmica terrestre, tais como em indústrias de telecomunicação via satélites.

Variando de acordo com o ciclo solar (de aproximadamente 11 anos), as manchas solares (do inglês *sunspots*) são caracterizadas pelo aparecimento de regiões com coloração mais escura sob a superfície de nossa estrela. Essas manchas, que podem durar meses, são relacionadas a outros tipos de atividade solar, e sua contagem auxilia na predição das condições de propagação de ondas curtas de rádio ou comunicações via satélite.

Por este motivo, o Observatório Real da Bélgica monitora, diariamente, desde 1749, o número de manchas na superfície solar e disponibiliza esses dados para pesquisas. Neste contexto, a série temporal histórica “*Sunspots Month*” traz, por mês, o número médio de manchas solares diárias (do inglês *Sunspots*). Ao analisar essa série (Figura 34a), nota-se a sazonalidade característica do fenômeno (aproximadamente 132 meses = 11 anos) e a diferença de amplitude (número de manchas) entre os ciclos. Como consequência o atrator (Figura 34b) apresenta formato cíclico, devido à sazonalidade dos dados, com tamanhos distintos relativos às diferenças de amplitudes. Cabe ressaltar que o atrator da Figura 34b foi construído em um espaço com seis dimensões e, por tal motivo, seu formato projetado torna-se, ainda, pior para visualização.

A presença de ciclos (sazonalidade) e padrão no comportamento da série temporal é uma característica que favorece a decomposição pelas técnicas **Spring** e suas variantes. Desta forma, é esperada uma melhoria nos valores preditos, após a etapa de decomposição.

7.1.3 Circuito de Chua

O Circuito de Chua (CHUA *et al.*, 1993) é um circuito eletrônico simples que apresenta um comportamento caótico na tensão. Com diversas aplicações reais, tais como em música (RODET, 1993) e criptografia (LIN; WANG, 2009), ele é composto por um circuito oscilatório linear convencional em paralelo com um outro circuito oscilatório não linear conhecido como diodo de Chua. Ao contrário de um oscilador eletrônico comum, ele produz uma onda oscilante não periódica, i.e., ondas que nunca se repetem, comportamento que pode ser facilmente simulado e modelado por equações não-lineares.

A fim de possibilitar a coleta dos dados e posterior estudo sobre a decomposição e predição do sinal, o Eng. Eletrônico Edson Emboaba de Oliveira projetou e construiu o circuito de Chua, seguindo as especificações descritas no trabalho original (CHUA *et al.*, 1993). Inicialmente, o projeto foi desenvolvido e simulado no software NI Multisim 14.0 (Figura 35a) e, após validação, construído de maneira física (Figura 35b).

Esse circuito foi ligado e sua diferença de potencial mensurada por meio de um osciloscópio. Devido à precisão do equipamento e interferências no sinal, os valores aferidos apresentam a adição de ruído branco, o que é evidente ao comparar seus resultados com o sinal simulado por equações diferenciais (CHUA, 1992). A presença de ruído resultante do processo de mensuração é uma das principais motivações para a decomposição e remoção de seu componente estocástico, objeto de estudo desta tese.

(a) Série temporal.

(b) Atrator da série temporal *Sunspots*. Cabe ressaltar que o atrator tem seis dimensões, mas apenas duas são representadas nesta figura.

Figura 34 – Série histórica mensal da média da quantidade de manchas solares desde 1749. Em (a), série temporal com os valores médios e, em (b), atrator da série imerso no espaço-fase.

Fonte: Elaborada pelo autor.

Com os valores coletados, e seguindo o mesmo processo adotado sobre os demais cenários, a série temporal (Figura 36a) foi imersa no espaço-fase (Figura 36b). O comportamento cíclico e com oscilações, que hora variam em valores positivos e hora em valores negativos, constrói um atrator com duas elipses que se conectam (comportamento que se assemelha ao do sistema de Lorenz (GRIGORENKO; GRIGORENKO, 2003)). As características inerentes do atrator de Chua, tais como transições suaves entre os estados, beneficiam as técnicas **Spring** e suas variações. Espera-se, portanto, que essas decomposições melhorem o resultados de predição dessa série temporal.

(a)

(b)

(c)

Figura 35 – Processo de simulação, construção e mensuração do Circuito de Chua. Em (a), a simulação do circuito no software NI Multisim 14.0, em (b), circuito construído e, em (c), a voltagem aferida por meio de um osciloscópio.

Fonte: Elaborada pelo autor.

(a) Série temporal.

(b) Atrator da série temporal Circuito de Chua. Cabe ressaltar que o atrator tem três dimensões, mas somente duas delas foram representadas nesta figura.

Figura 36 – Série temporal com a medição de voltagem do circuito de Chua. Em (a), série temporal com os valores da voltagem e, em (b), atrator da série imerso no espaço-fase.

Fonte: Elaborada pelo autor.

7.2 Metodologia dos Experimentos e Resultados

A previsão de séries temporais é um dos principais desafios da área de Aprendizado de Máquina porque a maioria dos algoritmos supervisionados pressupõe independência entre os valores (MELLO; PONTI, 2018). Essa restrição, na maioria das vezes, não pode ser garantida por uma série cujos valores tem uma dependência temporal. Esse cenário é ainda pior em séries determinísticas cujo valores são apenas explicados por uma combinação de valores anteriores (ALLIGOOD; SAUER; YORKE, 1996; KANTZ; SCHREIBER, 2004).

Neste contexto, diversos algoritmos foram desenvolvidos com o objetivo de prever os próximos valores de uma série temporal levando em consideração suas observações anteriores. Essa auto-regressão, por assim dizer, é significativamente afetada por ruídos que distorcem a posição dos valores passados, resultando em erros na predição de valores futuros.

Por este motivo, a decomposição de uma série temporal em seus respectivos componen-

tes determinístico e estocástico melhora o processo de predição, pois o ruído (estocástico) é separado do sinal que caracteriza o comportamento do fenômeno (determinístico). Ao executar as técnicas de predição sobre a série temporal sem ruído, não havendo erros expressivos nos valores anteriores, a auto-regressão projeta menores perdas e, conseqüentemente, obtém melhores resultados.

Para validar esta hipótese e comprovar a melhora esperada na etapa de predição, as três séries temporais reais, descritas neste capítulo, foram submetidas a dois cenários de teste: (i) a série original foi entregue a um algoritmo de predição, e (ii) a série original foi decomposta e o componente determinístico resultante foi entregue ao algoritmo de predição. Os resultados de ambos cenários foram então comparados aos valores aferidos da série e mensurados seus desvios por meios das medidas MAE e MDDL. Espera-se, portanto, que os resultados do cenário (ii) sejam melhores que a predição direta sobre os dados sem decomposição – cenário (i) – validando a hipótese deste trabalho.

Na etapa de predição foram utilizados dois algoritmos disponibilizados no pacote Tisean (HEGGER; KANTZ; SCHREIBER, 1999): (i) o polinomial e (ii) o funções de base radial. O primeiro, e mais simples, utiliza um polinômio de grau n para explicar o comportamento aferido dos valores da série. Esse polinômio é posteriormente utilizado para prever os próximos valores. O segundo algoritmo estima funções de base radial (do inglês *Radial Basis Functions* - RBF)(RANK, 2003) sobre os valores passados, a fim de estimar valores seguintes da série. Ambas estratégias necessitam de uma parametrização por parte do usuário que, a depender do conjunto de hiper-parâmetros escolhidos, pode convergir a resultados sub-ótimos. Para mitigar o efeito da parametrização e tornar a comparação mais justa, foi executada uma busca em grid (do inglês *grid search*) pela melhor configuração de parâmetros. A configuração ideal é tida como aquela que apresenta o menor valor para MAE e MDDL.

Para a decomposição, as três técnicas desenvolvidas no contexto deste trabalho foram utilizadas: (i) **Spring** (Seção 4.1), (ii) **Spring*** (Seção 4.3) e (iii) **Spring Time Domain** (Seção 4.2). Em um cenário em que tais técnicas não funcionem, o melhor valor obtido para os critérios de avaliação deverá ser maior ou igual ao cenário em que as técnicas de decomposição não foram utilizadas.

Os resultados de todos os testes estão consolidados na Figura 37. Cada par de coluna corresponde a um horizonte de tempo predito: 1, 5, 10, 25, 50, 75 e 100 observações futuras. Essas estão divididas entre as colunas para predição sem decomposição (cenário (i) – coluna P) e as colunas para decomposição e posterior predição (cenário (ii) – coluna D+P). Cada um desses cenários foi avaliado pelas medidas MAE e MDDL. Desta maneira, espera-se que os resultados presentes na coluna D+P tenham menores valores que os apresentados na coluna P, validando assim a hipótese desta tese e comprovando a utilidade das abordagens de decomposição propostas.

Cada um dos três conjuntos de dados reais (Sunspot, S&P500 e Circuito de Chua) está

exibido na Tabela da Figura 37, em grandes blocos. Esses estão divididos nos dois algoritmos de predição (Polinomial e funções de base radial (RBF)), combinados aos três algoritmos de decomposição (**Spring**, **Spring*** e **Spring Time Domain**). Para facilitar a análise dos resultados e validação da hipótese, todas as células da coluna D+P que têm menores valores que suas correspondentes na coluna P estão destacadas, em verde, por corroborarem com a hipótese. Ao contrário, todas as células da coluna P com valores menores são destacadas em amarelo. A hipótese está então validada se o número de células verde for superior ao número de amarelas.

Qtd de valores predito		Sunspot															
		1		5		10		25		50		75		100			
		P	D+P	P	D+P	P	D+P	P	D+P	P	D+P	P	D+P	P	D+P	P	D+P
Sunspot	Polinomial	MAE	Spring	0,679	0,064	13,120	10,166	19,142	16,123	13,645	14,895	19,251	17,305	27,434	20,900	26,255	22,623
		MAE	Spring*	0,679	0,679	13,120	13,120	19,142	19,142	13,645	13,645	19,251	19,251	27,434	27,434	26,255	26,255
		MAE	Spring TD	0,679	0,029	13,120	6,669	19,142	8,509	13,645	6,260	19,251	7,080	27,434	29,017	26,255	49,107
		MDDL	Spring	0,000	0,000	0,786	0,314	1,560	0,856	5,511	2,571	6,554	5,813	9,939	9,138	8,599	9,012
		MDDL	Spring*	0,000	0,000	0,786	0,786	1,560	1,560	5,511	5,511	6,554	6,554	9,939	9,939	8,599	8,599
		MDDL	Spring TD	0,000	0,000	0,786	0,387	1,560	0,711	5,511	1,097	6,554	3,132	9,939	4,589	8,599	19,174
	RBF	MAE	Spring	0,574	0,123	8,398	2,402	11,284	6,850	10,355	5,519	14,691	10,138	34,688	17,053	31,016	16,020
		MAE	Spring*	0,574	0,574	8,398	8,398	11,284	11,284	10,355	10,355	14,691	14,691	34,688	34,688	31,016	31,016
		MAE	Spring TD	0,574	14,464	8,398	10,889	11,284	11,132	10,355	10,793	14,691	11,874	34,688	31,552	31,016	29,363
		MDDL	Spring	0,000	0,000	0,376	0,348	0,847	0,536	1,121	0,695	3,224	3,055	7,046	6,301	8,060	6,369
		MDDL	Spring*	0,000	0,000	0,376	0,376	0,847	0,847	1,121	1,121	3,224	3,224	7,046	7,046	8,060	8,060
		MDDL	Spring TD	0,000	0,000	0,376	0,471	0,847	0,657	1,121	1,424	3,224	3,862	7,046	8,353	8,060	7,210
S&P 500	Polinomial	MAE	Spring	2,013	0,002	2,654	0,531	3,029	0,489	2,938	0,706	5,799	4,124	7,912	3,630	9,176	3,148
		MAE	Spring*	2,013	2,344	2,654	2,980	3,029	3,347	2,938	3,238	5,799	5,873	7,912	7,874	9,176	9,086
		MAE	Spring TD	2,013	2,061	2,654	2,799	3,029	3,314	2,938	3,676	5,799	5,270	7,912	5,527	9,176	5,392
		MDDL	Spring	0,000	0,000	0,786	0,421	2,419	0,555	7,827	1,993	8,155	4,725	12,923	6,358	19,515	5,957
		MDDL	Spring*	0,000	0,000	0,786	0,786	2,419	2,419	7,827	7,827	8,155	8,367	12,923	12,994	19,515	19,557
		MDDL	Spring TD	0,000	0,000	0,786	0,786	2,419	2,419	7,827	8,312	8,155	8,218	12,923	11,615	19,515	14,527
	RBF	MAE	Spring	0,039	0,000	0,814	0,484	1,097	0,465	1,459	0,672	5,585	2,759	5,203	2,876	4,155	2,343
		MAE	Spring*	0,039	0,199	0,814	0,817	1,097	1,142	1,459	1,432	5,585	5,575	5,203	6,857	4,155	7,413
		MAE	Spring TD	0,039	0,087	0,814	0,631	1,097	0,814	1,459	0,992	5,585	4,552	5,203	4,059	4,155	3,801
		MDDL	Spring	0,000	0,000	0,530	0,348	1,091	0,569	3,483	2,050	6,560	4,203	8,013	5,435	6,915	4,556
		MDDL	Spring*	0,000	0,000	0,530	0,530	1,091	1,332	3,483	3,655	6,560	5,589	8,013	11,580	6,915	18,516
		MDDL	Spring TD	0,000	0,000	0,530	0,530	1,091	0,954	3,483	2,386	6,560	2,998	8,013	6,768	6,915	5,990
Circuito de Chua	Polinomial	MAE	Spring	0,751	0,763	1,351	1,627	1,611	1,651	1,438	1,341	1,860	1,883	2,447	1,992	3,899	2,951
		MAE	Spring*	0,751	0,751	1,351	1,351	1,611	1,611	1,438	1,438	1,860	1,860	2,447	2,447	3,899	3,899
		MAE	Spring TD	0,751	0,043	1,351	1,013	1,611	1,256	1,438	1,221	1,860	1,771	2,447	2,155	3,899	5,678
		MDDL	Spring	0,000	0,000	0,000	0,309	0,474	0,420	0,805	0,471	1,467	1,011	1,846	1,173	2,091	1,358
		MDDL	Spring*	0,000	0,000	0,000	0,000	0,474	0,474	0,805	0,805	1,467	1,467	1,846	1,846	2,091	2,091
		MDDL	Spring TD	0,000	0,000	0,000	0,000	0,474	0,000	0,805	0,449	1,467	1,104	1,846	1,361	2,091	1,871
	RBF	MAE	Spring	1,831	0,006	3,600	2,689	8,399	2,652	6,195	4,497	7,487	6,591	9,409	9,193	16,203	17,182
		MAE	Spring*	1,831	1,831	3,600	3,600	8,399	8,399	6,195	6,195	7,487	7,487	9,409	9,409	16,203	16,203
		MAE	Spring TD	1,831	2,204	3,600	1,811	8,399	5,845	6,195	6,824	7,487	6,722	9,409	7,140	16,203	13,833
		MDDL	Spring	0,000	0,000	0,387	0,348	0,704	0,466	1,128	0,896	1,513	1,567	1,810	1,910	2,962	2,358
		MDDL	Spring*	0,000	0,000	0,387	0,387	0,704	0,704	1,128	1,128	1,513	1,513	1,810	1,810	2,962	2,962
		MDDL	Spring TD	0,000	0,000	0,387	0,348	0,704	0,603	1,128	1,204	1,513	1,695	1,810	1,861	2,962	2,605

Figura 37 – Comparação par-a-par das técnicas de decomposição. Percentual das vezes em que uma dada técnica (linha) supera o resultado das demais (colunas) com relação às medidas MAE e MDDL.

Fonte: Elaborada pelo autor.

Analisando os resultados, destacam-se as técnicas **Spring** e **Spring Time Domain** que, na maioria dos cenários, apresentaram melhores valores de MAE e MDDL que o cenário sem decomposição. Essa evidência se torna ainda mais forte à medida que a janela de predição aumenta. Este fato comprova a hipótese de que, ao decompor a série original, há um incremento

no desempenho relativo das predições. Por sua vez, ao comparar as técnicas entre si, a **Spring Time Domain** obteve melhores resultados para os dados de *Sunspots*. De fato, ao se analisar o atrator dessa série (Figura 34b), observam-se transições suaves entre os estados, o que está alinhado ao viés da técnica **Spring Time Domain**.

Devido ao cenário quase aleatório dos dados S&P 500, a técnica **Spring** destacou-se por eliminar ruídos e melhorar o processo de predição. Esses resultados não foram observados com as demais técnicas de decomposição. No circuito de Chua, as técnicas **Spring** e **Spring Time Domain** mais uma vez apresentaram resultados superiores ao cenário sem decomposição. Não há, no entanto, neste contexto, uma técnica de decomposição que se destaque: hora uma tem melhores resultados, hora outra. Chama a atenção os resultados obtidos pela técnica **Spring*** que, na maioria dos testes, obteve seu melhor resultado igual ao cenário sem decomposição. Este fato indica que a técnica não paramétrica não consegue remover suficientemente os ruídos, a ponto de melhorar a predição, portanto sua utilização não é aconselhada no contexto de previsões.

7.3 Considerações Finais

A presença de ruídos nas séries temporais dificulta o aprendizado e consequentemente a operação dos algoritmos de predição. Dessa maneira, a hipótese deste trabalho é que, ao decompor a série temporal em seus respectivos componentes determinísticos e estocásticos (este tipicamente associado a ruídos), há uma melhoria nos resultados de predição. Esta hipótese é particularmente verdadeira, caso a técnica de decomposição consiga separar ambos os vieses. Para demonstrar a eficácia das técnicas de decomposição desenvolvidas, bem como validar a hipótese supracitada, três bases de dados reais foram estruturadas, decompostas e tiveram seus valores preditos. Os resultados confirmam um bom desempenho das técnicas **Spring** e **Spring Time Domain**, bem como a validação da hipótese desta tese.

CONCLUSÕES

8.1 Conclusões Finais

A experiência prévia de um pesquisador faz com que ele/ela utilize a abordagem Estatística ou a baseada em Sistemas Dinâmicos para analisar e compreender séries temporais (RIOS, 2013). A utilização de apenas uma dessas abordagens leva, eventualmente, a erros de análise por desprezar influências estocásticas ou determinísticas. Faz-se, assim, necessária a decomposição de séries temporais para que cada componente seja tratado utilizando o ferramental mais adequado ao seu viés.

Diversas técnicas se propõem a tal decomposição, contudo elas impõem algum viés sobre o componente determinístico obtido, sendo o senoidal dominante (RIOS; MELLO, 2013). Essa imposição distorce relações temporais, destruindo possíveis recorrências e tendências originais da série. Esse problema torna evidente a necessidade de uma técnica capaz de decompor uma série temporal, preservando, contudo, o viés determinístico original da série. Esta limitação é contornada pelas técnicas de decomposição não lineares, mas elas impõem ao usuário a responsabilidade de sua sensível parametrização, além de dificuldades em lidar com diferenças de densidades nos estados presentes no espaço-fase.

Neste contexto, o objetivo desta tese foi o de projetar uma técnica de decomposição de séries temporais, preservando, contudo, o viés determinístico original da série. Além disto, foi de suma importância sua simples parametrização. Como consequência desta necessidade, foram desenvolvidas três abordagens: **Spring**, **Spring*** e **Spring TimeDomain**. Trabalhando diretamente sobre o atrator no espaço-fase, essas técnicas tentam remover a influência estocástica, filtrando o atrator da série temporal original.

Diversos experimentos foram efetuados para compreender e comprovar a eficácia dessas abordagens. Em um primeiro momento, a inspeção visual foi utilizada para compreender o viés imposto por tais técnicas durante o processo de decomposição. Além disso, pode-se observar

indícios de melhores resultados por parte das técnicas **Spring**, **Spring*** e **Spring TimeDomain** ao preservar o viés original da série temporal (Seção 6.1). Esses resultados foram corroborados por meio de uma análise quantitativa, baseada nas medidas MAE e MDDL (Seção 6.2). Com particular destaque, **Spring** produziu os melhores resultados para a maioria dos cenários.

Com o intuito de validar a hipótese deste trabalho, três series temporais reais foram mapeadas e utilizadas no processo de predição, são elas: (i) *Sunspots*, (ii) S&P500 e (iii) Circuito de Chua. Os experimentos realizados consistiram em verificar as medidas MAE e MDDL para o cenário de predição com e sem decomposição prévia. A hipótese estaria validada caso o cenário com decomposição produzisse menores valores de erro, fato confirmado pelo emprego da **Spring** e **Spring-TD**. Infelizmente, o mesmo não pode ser verificado para a técnica **Spring*** que, ao fim, demonstrou ter bom desempenho para decomposição, mas ainda insuficiente para auxiliar o processo de predição.

8.2 Publicações e Colaborações

Os diversos resultados obtidos ao longo do desenvolvimento desta tese de doutorado foram reportados nos seguintes artigos:

- DUARTE, FELIPE S. L. G.; RIOS, RICARDO A.; HRUSCHKA, EDUARDO R.; MELLO, RODRIGO F. de. Decomposing Time Series into Deterministic and Stochastic Influences: A Survey. Digital Signal Processing - Elsevier, 2019;
- DUARTE, FELIPE S. L. G.; RIOS, RICARDO A.; HRUSCHKA, EDUARDO R.; MELLO, RODRIGO F. de. Time series decomposition using spring system applied on phase spaces. In: 7th Brazilian Conference on Intelligent Systems, BRACIS 2018, São Paulo, Brazil, October 22-25, 2018. [s.n.], 2018. p.504–509;
- COSTA, FAUSTO G. DA ; DUARTE, FELIPE S.L.G. ; VALLIM, ROSANE M.M. ; MELLO, RODRIGO F. DE . Multidimensional surrogate stability to detect data stream concept drift. EXPERT SYSTEMS WITH APPLICATIONS, p. 15-29, 2017.

8.3 Trabalhos Futuros

Diversas oportunidades surgiram e foram mapeadas no desenvolvimento deste trabalho. Como explicado no início do Capítulo 4, seu foco é motivado por séries temporais em que o componente determinístico é parte dominante no sinal em análise. Assim, um estudo mais aprofundado sobre o funcionamento das técnicas de decomposição com relação a taxa sinal ruído da série temporal seria pertinente para entender o comportamento e as limitações de cada técnicas e, finalmente, ajudar o usuário a decidir sobre a melhor abordagem a depender da característica dos dados. Ainda, todo o estudo foi desenvolvido levando-se em consideração

que o ruído (componente estocástico) é somado ao componente determinístico formando a série temporal final. Cabe ressaltar que não há estudos conclusivos sobre o comportamento dessas técnicas em cenários em que os componentes são multiplicados, e não somados.

No que diz respeito ao desenvolvimento das técnicas, há indícios de que as abordagens desenvolvidas (**Spring**, **Spring*** e **Spring Time Domain**) podem ser melhoradas com estudos mais aprofundados sobre seus critérios de parada e correções durante as iterações. Também, outras abordagens para movimentar os estados e corrigir suas posições relativas no espaço-fase podem ser desenvolvidas e avaliadas. Ao longo deste processo, algumas abordagens foram testadas e obtiveram resultados promissores, porém aquém do esperado. Por exemplo, cabe mencionar que uma estratégia baseada na análise de componentes principais (PCA) foi projetada para determinar a direção e magnitude da movimentação de cada estado, contudo seus resultados, apesar de promissores, não foram equiparáveis ao estado da arte, logo essa abordagem não foi incluída neste documento final. Tentativas de correções durante o processo iterativo foram ainda avaliadas e, infelizmente não apresentaram bons resultados iniciais.

REFERÊNCIAS

ALLIGOOD, K. T.; SAUER, T. D.; YORKE, J. A. **Chaos: An Introduction to Dynamical Systems**. [S.l.]: Springer, 1996. Citado nas páginas 25, 26, 34, 39, 40, 89 e 106.

ARADHYE, G.; RAO, A. C. S.; MOHAMMED, M. D. M. A novel hybrid approach for time series data forecasting using moving average filter and ARIMA-SVM. In: **Advances in Intelligent Systems and Computing**. [S.l.]: Springer Singapore, 2018. p. 369–381. Citado na página 56.

AZCARRAGA, J.; SUAREZ, M. T. Predicting academic emotions based on brainwaves, mouse behaviour and personality profile. In: ANTHONY, P.; ISHIZUKA, M.; LUKOSE, D. (Ed.). **PRI-CAI 2012: Trends in Artificial Intelligence**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 728–733. ISBN 978-3-642-32695-0. Citado na página 56.

BENESTY, J.; CHEN, J.; HUANG, Y.; COHEN, I. Pearson correlation coefficient. In: **Noise reduction in speech processing**. [S.l.]: Springer, 2009. p. 1–4. Citado na página 85.

BERNARD, A. B.; JONES, C. I. Productivity across industries and countries: Time series theory and evidence. **The Review of Economics and Statistics**, JSTOR, v. 78, n. 1, p. 135, feb 1996. Citado na página 25.

BEVERIDGE, S.; NELSON, C. R. A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the ‘business cycle’. **Journal of Monetary Economics**, Elsevier BV, v. 7, n. 2, p. 151–174, jan 1981. Citado na página 56.

BISHOP, C. M. **Pattern Recognition and Machine Learning (Information Science and Statistics)**. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2006. ISBN 0387310738. Citado na página 73.

BOWERMAN, B. **Time series forecasting: Unified concepts and computer implementation**. [S.l.]: Duxbury Press, 1987. ISBN 9780871500700. Citado na página 38.

BOX, G. E. P.; JENKINS, G. **Time Series Analysis, Forecasting and Control**. 3º. ed. [S.l.]: Holden-Day, Incorporated, 1976. Citado nas páginas 25, 33, 34, 35, 37 e 57.

BOX, G. E. P.; PIERCE, D. A. Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. **Journal of the American Statistical Association**, Informa UK Limited, v. 65, n. 332, p. 1509–1526, dec 1970. Citado na página 38.

BRACEWELL, R. N. **The Fourier transform and its applications**. [S.l.]: McGraw-Hill New York, 1986. v. 31999. Citado nas páginas 51, 52 e 54.

BRANDT, M.; ROMANKIEWICZ, C.; SPIEKERMANN, R.; SAMIMI, C. Environmental change in time series - an interdisciplinary study in the sahel of mali and senegal. **Journal of Arid Environments**, Elsevier BV, v. 105, p. 52–63, jun 2014. Citado na página 25.

- BURNS, P. J. Robustness of the Ijung-box test and its rank equivalent. **SSRN Electronic Journal**, Elsevier BV, 2003. Disponível em: <<https://doi.org/10.2139/ssrn.443560>>. Citado na página 79.
- CANDAN, C.; KUTAY, M. A.; OZAKTAS, H. The discrete fractional fourier transform. **IEEE Transactions on Signal Processing**, Institute of Electrical and Electronics Engineers (IEEE), v. 48, n. 5, p. 1329–1337, maio 2000. Citado na página 52.
- CAO, J.; TAGLIASACCHI, A.; OLSON, M.; ZHANG, H.; SU, Z. Point cloud skeletons via laplacian based contraction. In: **Shape Modeling International Conference (SMI), 2010**. [S.l.]: IEEE, 2010. p. 187–197. Citado na página 71.
- CHUA, L. **The genesis of Chua's circuit**. [S.l.]: Electronics Research Laboratory, College of Engineering, 1992. Citado na página 103.
- CHUA, L. O.; ITOH, M.; KOCAREV, L.; ECKERT, K. Chaos synchronization in chua's circuit. **Journal of Circuits, Systems and Computers**, World Scientific Pub Co Pte Lt, v. 03, n. 01, p. 93–108, 1993. Citado na página 103.
- DEMSAR, J. Statistical comparisons of classifiers over multiple data sets. **Journal of Machine Learning Research**, JMLR.org, v. 7, p. 1–30, dez. 2006. ISSN 1532-4435. Citado na página 99.
- DEY, T. K.; SUN, J. Defining and computing curve-skeletons with medial geodesic function. In: **Proceedings of the Fourth Eurographics Symposium on Geometry Processing**. Aire-la-Ville, Switzerland, Switzerland: Eurographics Association, 2006. (SGP '06), p. 143–152. ISBN 3-905673-36-3. Citado na página 71.
- DUARTE, F. S.; RIOS, R. A.; HRUSCHKA, E. R.; MELLO, R. F. de. Decomposing time series into deterministic and stochastic influences: A survey. **Digital Signal Processing**, Elsevier BV, p. 102582, out. 2019. Disponível em: <<https://doi.org/10.1016/j.dsp.2019.102582>>. Citado nas páginas 51 e 83.
- DUARTE, F. S. L. G.; RIOS, R. A.; HRUSCHKA, E. R.; MELLO, R. F. d. Time series decomposition using spring system applied on phase spaces. In: **7th Brazilian Conference on Intelligent Systems, BRACIS 2018, São Paulo, Brazil, October 22-25, 2018**. [S.l.]: IEEE, 2018. p. 504–509. Citado nas páginas 56 e 72.
- ECKMANN, J.-P.; KAMPHORST, S. O.; RUELLE, D. Recurrence plots of dynamical systems. **EPL (Europhysics Letters)**, v. 4, n. 9, p. 973, 1987. Citado na página 63.
- FRASER, A. M.; SWINNEY, H. L. Independent coordinates for strange attractors from mutual information. **Physical review A**, APS, v. 33, n. 2, p. 1134, 1986. Citado nas páginas 43, 44 e 45.
- GOLUB, G.; KAHAN, W. Calculating the singular values and pseudo-inverse of a matrix. **Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis**, Society for Industrial & Applied Mathematics (SIAM), v. 2, n. 2, p. 205–224, jan 1965. Citado na página 58.
- GOLYANDINA, N. **Analysis of Time Series Structure: SSA and related techniques**. 1 edition. ed. [S.l.]: Chapman and Hall/CRC, 2001. (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). ISBN 9781420035841. Citado nas páginas 51 e 58.

GOMEZ, V.; MARAVALL, A. Programs tramo (time series regression with arima noise, missing observations, and outliers) and seats (signal extraction in arima time series). instructions for the user. **Documento de Trabajo**, v. 9628, 1996. Citado na página 56.

GRAPS, A. An introduction to wavelets. **IEEE Computing in Science and Engineering**, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 2, n. 2, p. 50–61, jun. 1995. ISSN 1070-9924. Citado nas páginas 51 e 54.

GRASSBERGER, P.; HEGGER, R.; KANTZ, H.; SCHAFFRATH, C.; SCHREIBER, T. On noise reduction methods for chaotic data. **Chaos: An Interdisciplinary Journal of Nonlinear Science**, AIP Publishing, v. 3, n. 2, p. 127–141, apr 1993. Citado nas páginas 52, 66 e 69.

GRIGORENKO, I.; GRIGORENKO, E. Chaotic dynamics of the fractional lorenz system. **Physical review letters**, APS, v. 91, n. 3, p. 034101, 2003. Citado na página 104.

GUZZO, F. **Detecção de mudanças de conceito em fluxos de dados por meio da análise de dependências temporais**. 115 p. Tese (Qualificação Doutorado) — Universidade de São Paulo, 2014. Citado na página 41.

HAAR, A. Zur theorie der orthogonalen funktionensysteme. **Mathematische Annalen**, Springer Science and Business Media LLC, v. 69, n. 3, p. 331–371, sep 1910. ISSN 1432-1807. Citado na página 55.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. [S.l.]: Springer, 2016. ISBN 0387848576. Citado na página 75.

HEGGER, R.; KANTZ, H.; SCHREIBER, T. Practical implementation of nonlinear time series methods: The tisean package. **CHAOS**, v. 9, n. 413, 1999. Citado nas páginas 26, 39, 52, 67 e 107.

HUANG, N. E.; SHEN, Z.; LONG, S. R.; WU, M. C.; SHIH, H. H.; ZHENG, Q.; YEN, N.-C.; TUNG, C. C.; LIU, H. H. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. **Proceedings of the Royal Society of London**, The Royal Society, v. 454, n. 1971, p. 903–995, mar 1998. Citado nas páginas 51, 54 e 59.

ISHII, R. P.; RIOS, R. A.; MELLO, R. F. Classification of time series generation processes using experimental tools: A survey and proposal of an automatic and systematic approach. **International Journal of Computational Science and Engineering**, Inderscience Publishers, Inderscience Publishers, Geneva, SWITZERLAND, v. 6, n. 4, p. 217–237, dez. 2011. ISSN 1742-7185. Citado nas páginas 33 e 34.

JIANG, J.; SHANG, P.; ZHANG, Z.; LI, X. The multi-scale high-order statistical moments of financial time series. **Physica A: Statistical Mechanics and its Applications**, Elsevier BV, v. 512, p. 474–488, dec 2018. Citado na página 35.

KANTZ, H.; SCHREIBER, T. **Nonlinear Time Series Analysis**. [S.l.]: Cambridge University Press, 2004. (Cambridge nonlinear science series). ISBN 9780521529020. Citado nas páginas 25, 33, 34, 43, 47, 51, 89 e 106.

KENNEL, M. B.; BROWN, R.; ABARBANEL, H. D. Determining embedding dimension for phase-space reconstruction using a geometrical construction. **Physical review A**, APS, v. 45, n. 6, p. 3403, 1992. Citado nas páginas 26, 43, 47 e 48.

- LAU, K.-M.; WENG, H. Climate signal detection using wavelet transform: How to make a time series sing. **Bulletin of the American Meteorological Society**, American Meteorological Society, v. 76, n. 12, p. 2391–2402, dec 1995. Citado na página 25.
- LI, J.; TENG, Z.; WANG, Y.; YOU, S.; LIU, Y.; YAO, W. A fast power grid frequency estimation approach using frequency-shift filtering. **IEEE Transactions on Power Systems**, Institute of Electrical and Electronics Engineers (IEEE), v. 34, n. 3, p. 2461–2464, may 2019. Citado na página 57.
- LI, W. Mutual information functions versus correlation functions. **Journal of Statistical Physics**, Springer Science and Business Media LLC, v. 60, n. 5-6, p. 823–837, sep 1990. ISSN 0022-4715. Citado na página 63.
- LIN, Z.-h.; WANG, H.-x. Image encryption based on chaos with PWL memristor in chua's circuit. In: **2009 International Conference on Communications, Circuits and Systems**. IEEE, 2009. Disponível em: <<https://doi.org/10.1109/icccas.2009.5250354>>. Citado na página 103.
- LJUNG, G. M.; BOX, G. E. P. On a measure of lack of fit in time series models. **Biometrika**, Oxford University Press (OUP), v. 65, n. 2, p. 297–303, aug 1978. Disponível em: <<https://doi.org/10.1093/biomet/65.2.297>>. Citado nas páginas 38 e 78.
- LU, H.; PLATANIOTIS, K. N.; VENETSANOPOULOS, A. **Multilinear Subspace Learning: Dimensionality Reduction of Multidimensional Data (Chapman & Hall/CRC Machine Learning & Pattern Recognition)**. [S.l.]: Chapman and Hall/CRC, 2013. ISBN 1439857245. Citado na página 58.
- MELLO, R. F.; PONTI, M. A. **Machine Learning: A Practical Approach on the Statistical Learning Theory**. [S.l.]: Springer, 2018. ISBN 3319949888. Citado na página 106.
- MENDELSON, B. **Introduction to Topology**. [S.l.]: Dover Publications, 2012. (Dover Books on Mathematics). ISBN 9780486135090. Citado nas páginas 56 e 69.
- MFOUMOU, G. S.; KENMOE, G. D.; KOFANE, T. Computational algorithms of time series for stick-slip dynamics and time-delayed feedback control of chaos for a class of discontinuous friction systems. **Mechanical Systems and Signal Processing**, Elsevier BV, v. 119, p. 399–419, mar 2019. Citado na página 35.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to Time Series Analysis and Forecasting**. [S.l.]: Wiley, 2015. (Wiley Series in Probability and Statistics). ISBN 978-0471653974. Citado na página 25.
- MORETTIN, P. A.; TOLOI, C. **Análise de series temporais**. [S.l.]: Edgard Blucher, 2006. (ABE - Projeto Fisher). ISBN 9788521203896. Citado nas páginas 25, 33, 35 e 38.
- MORICE, C. P.; KENNEDY, J. J.; RAYNER, N. A.; JONES, P. D. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The hadcrut4 data set. **Journal of Geophysical Research: Atmospheres**, American Geophysical Union (AGU), v. 117, n. D8, apr 2012. ISSN 2156-2202. Citado na página 36.
- MULLER, M. **Information retrieval for music and motion**. [S.l.]: Springer, 2007. v. 2. Citado na página 86.

OGNIEWICZ, R. L. **Automatic Medial Axis Pruning by Mapping Characteristics of Boundaries Evolving under the Euclidean Geometric Heat Flow onto Voronoi Skeletons**. [S.l.], 1995. Citado na página 71.

OTT, E.; SAUER, T.; YORKE, J. A. **Coping with chaos. Analysis of chaotic data and the exploitation of chaotic systems**. [S.l.]: John Wiley & Sons, 1994. v. 1. Citado nas páginas 42 e 43.

PAGLIOSA, L. d. C.; MELLO, R. F. d. Applying a kernel function on time-dependent data to provide supervised-learning guarantees. **Expert Systems with Applications**, Pergamon Press, Inc., Tarrytown, NY, USA, v. 71, n. C, p. 216–229, 2017. ISSN 0957-4174. Citado na página 89.

Palis Junior, J. **Geometry and topology : proceedings of the school held at the Instituto de Matematica Pura e Aplicada CNPq, Rio de Janeiro, July 1976**. Berlin New York: Springer-Verlag, 1977. ISBN 978-3-540-37301-8. Citado na página 66.

PAN, V. **Structured Matrices and Polynomials: Unified Superfast Algorithms**. [S.l.]: Boston, Birkhauser, 2012. ISBN 9781461201298. Citado na página 58.

PERCIVAL, D. **Wavelet methods for time series analysis**. Cambridge New York: Cambridge University Press, 2000. ISBN 978-0521685085. Citado na página 55.

RAMASWAMY, P. Utilizing gamma band to improve mental task based brain-computer interface design. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, Institute of Electrical and Electronics Engineers (IEEE), v. 14, n. 3, p. 299–303, sep 2006. Citado na página 54.

RANK, E. Application of bayesian trained rbf networks to nonlinear time-series modeling. **Signal Processing**, Elsevier BV, v. 83, n. 7, p. 1393–1410, 2003. Disponível em: <[https://doi.org/10.1016/s0165-1684\(03\)00088-4](https://doi.org/10.1016/s0165-1684(03)00088-4)>. Citado na página 107.

RIOS, R. A. **Improving time series modeling by decomposing and analysing stochastic and deterministic influences**. Tese (Doutorado) — University of São Paulo, 10 2013. Citado nas páginas 26, 34, 35, 51, 60, 61, 63, 64 e 111.

RIOS, R. A.; MELLO, R. F. Improving time series modeling by decomposing and analyzing stochastic and deterministic influences. **Signal Processing**, Elsevier BV, v. 93, n. 11, p. 3001 – 3013, nov 2013. ISSN 0165-1684. Citado nas páginas 51, 52, 87 e 111.

RIOS, R. A.; PARROTT, L.; LANGE, H.; MELLO, R. F. d. Estimating determinism rates to detect patterns in geospatial datasets. **Remote Sensing of Environment**, Elsevier BV, v. 156, p. 11 – 20, jan 2015. ISSN 0034-4257. Citado na página 54.

RODET, X. Sound and music from chua's circuit. **Journal of Circuits, Systems and Computers**, World Scientific Pub Co Pte Lt, v. 03, n. 01, p. 49–61, 1993. Citado na página 103.

ROUSSEEUW, P.; PERROTTA, D.; RIANI, M.; HUBERT, M. Robust monitoring of time series with application to fraud detection. **Econometrics and Statistics**, Elsevier BV, v. 9, p. 108–121, jan 2019. Citado na página 35.

RUBIK, B. Neurofeedback-enhanced gamma brainwaves from the prefrontal cortical region of meditators and non-meditators and associated subjective experiences. **The Journal of Alternative and Complementary Medicine** Vol. 17, No. 2 Original Articles, Mary Ann Liebert Inc, v. 17, n. 2, p. 109–115, Feb. 2011. ISSN 1557-7708. Citado na página 56.

- SABZIKAR, F.; MCLEOD, A. I.; MEERSCHAERT, M. M. Parameter estimation for ARTFIMA time series. **Journal of Statistical Planning and Inference**, Elsevier BV, v. 200, p. 129–145, may 2019. Citado na página 35.
- SAID, S. E.; DICKEY, D. A. Testing for unit roots in autoregressive-moving average models of unknown order. **Biometrika**, Oxford University Press (OUP), v. 71, n. 3, p. 599–607, 1984. Citado na página 56.
- SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. **Analytical Chemistry**, American Chemical Society (ACS), v. 36, n. 8, p. 1627–1639, jul 1964. Citado nas páginas 56 e 57.
- SCHREIBER, T. Extremely simple nonlinear noise-reduction method. **Physical Review E**, American Physical Society (APS), v. 47, n. 4, p. 2401–2404, apr 1993. Disponível em: <<https://doi.org/10.1103/physreve.47.2401>>. Citado nas páginas 52, 66, 69 e 73.
- SOKLAKOV, A. Occams razor as a formal basis for a physical theory. **Foundations of Physics Letters**, Springer Nature, v. 15, n. 2, p. 107–135, 2002. ISSN 0894-9875. Citado na página 37.
- STOCCHI, M.; MARCHESI, M. Fast wavelet transform assisted predictors of streaming time series. **Digital Signal Processing**, Elsevier BV, v. 77, p. 5–12, jun 2018. Citado na página 54.
- STRANG, G. **Introduction to Linear Algebra**. [S.l.]: Wellesley-Cambridge Press, 2003. ISBN 9780961408893. Citado na página 58.
- TAGLIASACCHI, A.; ZHANG, H.; COHEN-OR, D. Curve skeleton extraction from incomplete point cloud. **ACM Transactions on Graphics**, Association for Computing Machinery (ACM), v. 28, n. 3, p. 1, jul 2009. Citado nas páginas 66 e 71.
- TAKENS, F. Detecting strange attractors in turbulence. In: **Dynamical systems and turbulence**. [S.l.]: Springer, 1981. p. 366–381. Citado nas páginas 26, 35, 39, 42, 43, 44, 52, 63, 66, 69, 72 e 87.
- TIPLER, P. A. **Física Para Cientistas e Engenheiros: Mecânica, Oscilações e Ondas, Termodinâmica - Vol.1**. [S.l.]: LTC, 2009. ISBN 8521617100. Citado na página 74.
- TORMENE, P.; GIORGINO, T.; QUAGLINI, S.; STEFANELLI, M. Matching incomplete time series with dynamic time warping: An algorithm and an application to post-stroke rehabilitation. **Artificial Intelligence in Medicine**, Elsevier Science Publishers Ltd., Essex, UK, v. 45, n. 1, p. 11–34, 2009. ISSN 0933-3657. Citado nas páginas 86 e 87.
- ZHANG, Y.; ZOU, M.; YANG, S.; LIN, Y.; YE, C. Colder-winter monthly-temperature forecasting in general trend of global warming via multiple sine functions decomposition. In: **2017 36th Chinese Control Conference (CCC)**. [S.l.]: IEEE, 2017. p. 4551–4555. Citado na página 89.