CEFET/RJ
DIRETORIA DE PESQUISA E PÓS-GRADUAÇÃO

TRANSFORMAÇÃO RESILIENTE PARA DETECÇÃO DE ANOMALIAS EM SÉRIES TEMPORAIS HETEROSCEDÁSTICAS

Laís Baroni Villet

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Produção e Sistemas, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ, como parte dos requisitos necessários à obtenção do título de doutor.

Orientadores:
Eduardo Soares Ogasawara
Rafaelli de Carvalho Coutinho

Rio de Janeiro

29-01-2026

Transformação Resiliente para Detecção de Anomalias em Séries Temporais Heteroscedásticas

Tese de Doutorado em Engenharia de Produção e Sistemas, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, CEFET/RJ.

Laís Baroni Villet

Aprovada por:

Presidente, Prof. Eduardo Soares Ogasawara, D.Sc. (orientador)

Rafaelli de Carvalho Coutinho, D.Sc. (coorientadora)

Gustavo Paiva Guedes e Silva, D.Sc.

Fábio André Machado Porto, D.Sc.

Carlos Eduardo Ribeiro de Mello, D.Sc.

Eduardo Fonseca Mendes, D.Sc.

Rio de Janeiro

29-01-2026

Ficha catalográfica elaborada pela Biblioteca Central do CEFET/RJ

V748 Villet, Laís Baroni
Transformação resiliente para detecção de anomalias em séries temporais heteroscedásticas / Laís Baroni Villet. — 2026.
64f. : il. color. , enc.

Tese (Doutorado) Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, 2026.
Bibliografia : f. 54-64
Orientador: Eduardo Soares Ogasawara
Coorientador: Rafaelli de Carvalho Coutinho

1. Análise de séries temporais - Anomalias . 2. Evento – Detecção. 3. Volatilidade. 4. Análise de variância. I. Ogasawara, Eduardo Soares. (Orient.). II. Coutinho, Rafaelli de Carvalho. (Coorient.). III. Título.

CDD 519.55

Elaborada pela bibliotecária Tania Mello – CRB/7 nº 5507/04

RESUMO

Transformação Resiliente para Detecção de Anomalias em Séries Temporais Heteroscedásticas

Laís Baroni Villet

Orientadores:
Eduardo Soares Ogasawara
Rafaelli de Carvalho Coutinho

Resumo da Tese submetida ao Programa de Pós-graduação em Engenharia de Produção e Sistemas do Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ como parte dos requisitos necessários à obtenção do título de doutor.

Séries temporais heteroscedásticas apresentam variância não constante ao longo do tempo, o que prejudica a eficácia dos métodos clássicos de detecção de anomalias baseados em limiares fixos. Em particular, essa característica afeta a detecção de anomalias contextuais, cuja detecção depende da dispersão local. Para abordar esse desafio, este trabalho propõe uma técnica de transformação denominada Transformação Resiliente (RT), que modifica a série original por meio da decomposição empírica de modos, da seleção de componentes de alta frequência com base na rugosidade, da diferenciação e da normalização pela dispersão local. Essa transformação equaliza a variância e destaca as anomalias. Avaliações utilizando o conjunto de dados Yahoo! S5, com o framework Harbinger, demonstram que a RT melhora o desempenho de métodos clássicos e pode ser combinada com limiarização para formar um método de detecção autônomo, denominado Detector de Anomalias por Transformação Resiliente (DATR).

Palavras-chave:
Séries temporais, transformação, detecção de eventos, anomalias, heteroscedasticidade, volatilidade, decomposição do modo empírico

Rio de Janeiro

29-01-2026

ABSTRACT

Transformação Resiliente para Detecção de Anomalias em Séries Temporais Heteroscedásticas

Laís Baroni Villet

Advisors:
Eduardo Soares Ogasawara
Rafaelli de Carvalho Coutinho

Abstract of tesis submitted to Programa de Pós-graduação em Engenharia de Produção e Sistemas - Centro Federal de Educação Tecnológica Celso Suckow da Fonseca CEFET/RJ as partial fulfillment of the requirements for the degree of doctor.

Heteroskedastic time series exhibit non-constant variance over time, which compromises the effectiveness of classical anomaly detection methods based on fixed thresholds. In particular, this characteristic affects the detection of contextual anomalies, whose detection relies on local dispersion. To address this challenge, this paper proposes a transformation technique called Resilient Transformation (RT), which modifies the original series through empirical mode decomposition, roughness-based high-frequency component selection, differencing, and local dispersion normalization. This transformation equalizes variance and highlights anomalies. Evaluations using the Yahoo! S5 dataset, within the Harbinger framework, show that RT improves the performance of classical methods and can be combined with thresholding to form a standalone detection method called Resilient Transformation Anomaly Detector (RTAD).

Key-words:
Time series, transformation, event detection, anomaly detection, heteroskedasticity, volatility, empirical mode decomposition

Rio de Janeiro

29-01-2026

Sumário

I Introdução **12**

II Contextualização **16**

II.1 Séries temporais 16

II.1.1 Normalização 17

II.1.2 Diferenciação 17

II.1.3 Janelas deslizantes 18

II.1.4 Heteroscedasticidade 18

II.1.5 Rugosidade 20

II.1.6 Curvatura máxima 21

II.2 Decomposição de séries temporais 22

II.3 Detecção de eventos 23

II.3.1 Anomalias 25

II.4 Métricas de avaliação 26

II.4.1 Testes Estatísticos e Tamanho do Efeito 27

III Trabalhos Relacionados **29**

IV Metodologia **35**

IV.1 Transformação Resiliente 35

IV.2 Detecção de Anomalias com RTAD 40

IV.3 Implementação 41

IV.4 Exemplo ilustrativo 42

V Avaliação Experimental **46**

V.1 Configuração Experimental 46

V.1.1 Conjunto de Dados 46

V.1.2 Medidas de Avaliação 47

V.1.3 Modelos de Linha de Base 48

V.1.4 Otimização de Hiperparâmetros 50

V.1.5 Validação Estatística 50
V.2 Experimentos 51
V.2.1 Eficácia da RT 51
V.2.2 Teste de Ablação: Análise no Conjunto de Séries Homoscedásticas 54
V.2.3 Detector de Anomalias por Transformação Resiliente (RTAD) 55
V.2.4 Sensibilidade do RTAD à Medida de Dispersão 57
V.2.5 Sensibilidade ao Critério de Limiar 59
V.3 Discussão 59

VI Conclusões **62**

Referências **65**

Lista de Figuras

II.1 Decomposição CEEMD. 24

IV.1 Fluxograma da Transformação Resiliente (RT) 36
IV.2 Valores de rugosidade das IMF. 38
IV.3 Séries a) δ_t, b) ψ_t e c) α_t. 40
IV.4 Série temporal Y_t de entrada do exemplo ilustrativo. 42
IV.5 IMFs geradas pela decomposição CEEMD da série Y_t. 43
IV.6 Componente residual gerado pela decomposição CEEMD da série Y_t. 43
IV.7 Valores de rugosidade das IMFs e ponto de seleção d. 43
IV.8 Série σ_t resultante da composição das IMFs selecionadas. 43
IV.9 Série δ_t no exemplo ilustrativo. 44
IV.10 Série ξ_t (série Y_t com a tendência removida). 44
IV.11 Série ψ_t de dispersão local no exemplo ilustrativo. 44
IV.12 Série α_t final transformada no exemplo ilustrativo. 44
IV.13 Detecção de anomalias na série α_t do exemplo. 45

V.1 Exemplo de séries usadas na avaliação experimental. 47
V.2 Comparação do desempenho de diferentes modelos: (a) F1, (b) revocação e (c) precisão, com e sem aplicação da RT. Cada ponto representa uma série temporal; a linha tracejada indica igualdade de desempenho, e o 'X' indica a média dos resultados para cada modelo. 52
V.3 Exemplo comparativo da detecção de anomalias com e sem aplicação da RT para os métodos FBIAD, ARIMA e GARCH em uma série temporal selecionada. O painel superior exibe a série temporal original. Os painéis inferiores detalham os resultados para cada configuração. Anomalias corretamente detectadas (Verdadeiros Positivos) são indicadas em verde, anomalias não detectadas (Falsos Negativos) em azul, e falsas detecções (Falsos Positivos) em vermelho. 54
V.4 Boxplots comparando F1, precisão e revocação de métodos de detecção para séries temporais heteroscedásticas. O método RTAD é destacado em cinza claro. 56

V.5 Tamanho do efeito para cada método e métrica. Valores sublinhados indicam diferenças estatisticamente significativas; asteriscos indicam casos em que o valor mediano da métrica para RTAD é menor do que o do método correspondente. 57
V.6 F1 do RTAD no conjunto de dados heteroscedástico usando diferentes medidas de dispersão para cálculo da volatilidade. 58
V.7 Métricas de avaliação no conjunto de dados heteroscedástico utilizando diferentes estratégias de limiarização. 59

Lista de Tabelas

III.1 Resumo comparativo dos trabalhos mais semelhantes à RT/RTAD. 34

IV.1 Símbolos utilizados na RT e no RTAD. 36

IV.2 Parâmetros opcionais da RT no contexto do RTAD. 42

V.1 Hiperparâmetros variados na otimização dos modelos de aprendizado de máquina. 50

V.2 Análise estatística comparativa de RT para métodos de detecção de anomalias em séries temporais heteroscedásticas. 53

V.3 Análise estatística comparativa da RT para métodos de detecção de anomalias em séries temporais homoscedásticas. 55

V.4 Descrição e justificativa das medidas de dispersão consideradas. 58

V.5 Redução percentual de séries heteroscedásticas por medida de dispersão, com base em 74 séries no conjunto de dados. 58

Lista de Acrônimos

ARCH	*Autoregressive Conditional Heteroskedasticity.*
CEEMD	*Complete Ensemble Empirical Mode Decomposition.*
EEMD	*Ensemble Empirical Mode Decomposition.*
EGARCH	*Exponencial Generalized Autoegressive Conditional Heteroskedasticity.*
EMD	*Empirical Mode Decomposition.*
FN	*falso negativo.*
FP	*falso positivo.*
GARCH	*Generalized Autoegressive Conditional Heteroskedasticity.*
GJR-GARCH	*Glosten-Jagannathan-Runkle Generalized Autoegressive Conditional Heteroskedasticity.*
IMF	*Intrinsic Mode Function.*
NAGARCH	*Nonlinear Asymmetric Generalized Autoegressive Conditional Heteroskedasticity.*
QGARCH	*Quadratic Generalized Autoegressive Conditional Heteroskedasticity.*
RT	*Resilient Transformation.*
RTAD	*Resilient Transformation Anomaly Detector.*
VN	*verdadeiro negativo.*
VP	*verdadeiro positivo.*

Capítulo I Introdução

A detecção de eventos em séries temporais constitui uma área de pesquisa consolidada e de grande relevância, com vastas aplicações em domínios científicos e práticos, como finanças, saúde, segurança cibernética, processos industriais e previsão ambiental [Braei and Wagner, 2020; Zou et al., 2023]. Séries temporais são sequências de observações ordenadas no tempo e registradas em intervalos regulares, formalmente expressas como $Y = \langle y_1, y_2, \ldots, y_n \rangle$, onde y_t representa o valor observado no tempo t. A análise desses dados permite a detecção de padrões, a previsão de comportamentos futuros e a detecção de eventos que se desviam do comportamento esperado do sistema observado [Wenig et al., 2024; Zhao et al., 2020].

Entre os eventos de interesse, destacam-se as anomalias, definidas como observações ou padrões que divergem significativamente do restante da série [Ogasawara et al., 2025b]. As anomalias podem ser classificadas como pontuais ou sequenciais, e também como globais ou contextuais, dependendo de sua relação com o comportamento local da série. As anomalias contextuais representam um desafio adicional, pois sua natureza anômala depende do regime de variabilidade em que ocorrem. A detecção dessas anomalias exige que o processo considere explicitamente o contexto estatístico local, que frequentemente é caracterizado e modelado por meio de distribuições condicionais, a fim de determinar a probabilidade de uma observação sob as condições prevalecentes [Zou et al., 2023; Wenig et al., 2024].

Esse desafio torna-se ainda mais pronunciado em séries temporais que exibem heteroscedasticidade condicional, um fenômeno onde a variância condicional (ou volatilidade) da série flutua ao longo do tempo. Nesses casos, flutuações locais na amplitude dos dados distorcem medidas tradicionais de dispersão e afetam diretamente a sensibilidade de detectores baseados em limiares fixos ou em premissas de homoscedasticidade. A desconsideração da heteroscedasticidade compromete a eficácia dos métodos de detecção, gerando estimativas enviesadas e alarmes incorretos. Por exemplo, a falha em identificar uma anomalia genuína, contextualmente significativa em períodos de baixa variância, pode acarretar perdas substanciais. Inversamente, a detecção de um evento como anômalo em um período de alta variância, onde tal comportamento é esperado, pode induzir decisões equivocadas, elevando a taxa de falsos positivos. Adicionalmente, métodos tradicionais dependem frequentemente de parametrizações complexas que não são robustas a dados não estacionários com variabilidade estrutural, o que dificulta sua aplicação em contextos reais e dinâmicos [Ogasawara

et al., 2025b; Zou et al., 2023]. Essa problemática sublinha uma lacuna significativa na literatura atual: a ausência de métodos de detecção resilientes e eficientes, capazes de lidar com a variabilidade e a complexidade inerentes às séries temporais heteroscedásticas sem impor premissas simplificadoras, como estacionariedade ou homoscedasticidade.

Ao longo deste trabalho, o termo heteroscedasticidade será empregado para se referir especificamente à heteroscedasticidade condicional. Enquanto a heteroscedasticidade incondicional implica que a variância global do processo estocástico muda ao longo do tempo (violando a estacionariedade de segunda ordem), a heteroscedasticidade condicional descreve a situação onde a variância de uma série temporal em um dado momento é condicionada por informações passadas ou por variáveis contextuais. Em muitos contextos de séries temporais a variância incondicional pode ser constante (caracterizando um processo estacionário de segunda ordem), mas a variância condicional exibe flutuações temporais. Consequentemente, discussões sobre 'variância variável' ou 'volatilidade' em capítulos subsequentes referem-se sempre à natureza condicional da variabilidade observada.

A prevalência da heteroscedasticidade é notória em diversos domínios críticos. Em dados dos mercados financeiros, por exemplo, os retornos diários de ações são intrinsecamente heteroscedásticos, com períodos alternados de alta e baixa variância condicional. Analogamente, em dados epidemiológicos, a heteroscedasticidade pode ser observada na flutuação do número de novos casos de uma doença ao longo do tempo, reflexo de fatores como a sazonalidade, eventos de superespalhamento ou o impacto de medidas de contenção. Nesses contextos, a detecção de anomalias pode revelar padrões importantes, como negociações fraudulentas ou a emergência precoce de surtos epidemiológicos atípicos, além de apoiar a tomada de decisão e a gestão de riscos.

Nesse contexto, a decomposição empírica surge como uma alternativa promissora para o tratamento de séries temporais complexas. Técnicas como a *Empirical Mode Decomposition* (EMD) e suas variantes, incluindo a *Ensemble Empirical Mode Decomposition* (EEMD) e a *Complete Ensemble Empirical Mode Decomposition* (CEEMD), permitem representar a série como a soma de componentes oscilatórios, conhecidos como *Intrinsic Mode Function* (IMF)s, e de um componente residual de baixa frequência associado a tendências de longo prazo [Huang et al., 1998; Souza et al., 2024]. Estas abordagens são particularmente úteis por serem adaptativas e orientadas pelos dados, sem a necessidade de premissas a priori sobre a estacionariedade ou linearidade da série.

A intuição fundamental que guia a presente abordagem reside em uma inversão conceitual da metodologia tradicional para detecção de anomalias. Enquanto métodos clássicos frequentemente se dedicam a modelar o comportamento esperado da série temporal para, em seguida, identificar desvios (resíduos) entre os valores observados e o modelo, a presente abordagem adota um caminho distinto. Em vez de primeiramente modelar a série, ela é concebida para transformar diretamente os dados de forma a realçar seu componente residual intrínseco — ou seja, os aspectos que, em um

contexto tradicional, seriam considerados ruído ou flutuações inesperadas que sinalizam anomalias.

Essa abordagem inovadora é materializada por um processo de transformação em duas etapas principais. Primeiramente, são empregadas técnicas que permitem a separação dos componentes da série temporal, focando naqueles que representam flutuações rápidas e irregulares, frequentemente associadas a eventos anômalos. Em seguida, busca-se equalizar a variabilidade local da série, um passo crucial para atenuar os efeitos da heteroscedasticidade e amplificar a visibilidade de comportamentos abruptos e transições significativas.

Com base nessa perspectiva inovadora, esta Tese propõe a *Resilient Transformation* (RT) e o *Resilient Transformation Anomaly Detector* (RTAD). A RT é uma metodologia concebida para transformar séries temporais heteroscedásticas, acentuando anomalias contextuais. Essa transformação se fundamenta na decomposição da série por meio da CEEMD para isolar seus componentes de alta frequência e curta duração (as *Intrinsic Mode Function* (IMF)s mais rugosas), seguida por uma normalização local da variância que pondera as primeiras diferenças pela volatilidade instantânea. O resultado é uma nova série transformada que preserva anomalias relevantes e apresenta amplitude homogênea ao longo do tempo. No âmbito do RTAD, as detecções baseiam-se na aplicação da RT seguida de uma decisão por limiar de dispersão estatística.

A tese central defendida neste trabalho é que a aplicação da RT a séries heteroscedásticas converte anomalias contextuais em anomalias globalmente detectáveis, possibilitando sua detecção por métodos clássicos de detecção de anomalias. A questão de pesquisa que guia este estudo é: *"É possível aprimorar a detecção de anomalias em séries temporais heteroscedásticas por meio de uma transformação que equalize a dispersão local?"*

Os objetivos deste trabalho são: (i) propor e formalizar a RT, descrevendo seus passos e a sua fundamentação teórica; (ii) avaliar o impacto da RT na capacidade de detecção de métodos clássicos de detecção de anomalias; (iii) investigar o desempenho da RT combinada com um critério de decisão simples como base de um método de detecção autônomo (RTAD); e (iv) realizar análises de sensibilidade em relação ao cálculo da dispersão estatística na RT e ao critério de limiar utilizado nas decisões do RTAD.

A principal contribuição desta Tese é metodológica: apresenta uma estratégia de transformação que aborda diretamente a heteroscedasticidade das séries temporais sem depender de modelos paramétricos complexos ou de ajustes específicos do domínio. A metodologia proposta não requer parametrizações complexas ou ajustes específicos para cada conjunto de dados. Em vez disso, opera com um conjunto fixo de parâmetros intrínsecos à sua concepção, projetados para atender eficientemente ao processo de detecção independentemente das características da série temporal de interesse. Essa abordagem reduz significativamente a subjetividade e a necessidade de calibração empírica na definição de parâmetros. Consequentemente, o método torna-se mais resiliente e confiável em dife-

rentes contextos e conjuntos de dados, o que acelera sua implementação e favorece a adoção prática.

O desempenho da metodologia proposta foi avaliado e comparado a métodos conhecidos na literatura, utilizando o conjunto de séries temporais com eventos rotulados do *Yahoo! Webscope dataset*. Esse conjunto de dados foi selecionado devido à sua heterogeneidade, que inclui séries sintéticas com variadas características de comportamento, como tendências, sazonalidade e distintos regimes de variância. Isso permitiu uma validação robusta da aplicabilidade geral da RT e do RTAD. Para a avaliação experimental, foram utilizadas as métricas F1-score, precisão e revocação, além de testes de significância estatística e de tamanho de efeito. Os resultados demonstraram que a RT melhora o desempenho de diversos métodos clássicos de detecção em séries heteroscedásticas. Além disso, quando combinada com um critério de decisão, o RTAD demonstrou desempenho competitivo com simplicidade operacional e estabilidade. Análises de sensibilidade indicaram que o uso do desvio padrão como medida de dispersão e o limiar gaussiano para tomada de decisão produz os melhores resultados, reforçando as escolhas conceituais da RT e do RTAD. Esses achados confirmam a utilidade da RT tanto como componente auxiliar quanto como método independente de detecção de anomalias.

Esta Tese está dividida em cinco capítulos. O Capítulo 2 apresenta os conceitos fundamentais que situam o presente trabalho. O Capítulo 3 revisa a literatura relevante. O Capítulo 4 detalha a metodologia proposta, incluindo a formalização da RT e os fundamentos do RTAD. O Capítulo 5 apresenta a avaliação experimental, incluindo os resultados da aplicação da RT com métodos clássicos de detecção, a análise de desempenho do RTAD e as análises de sensibilidade. Finalmente, o Capítulo 6 conclui os resultados obtidos e discute suas implicações e limitações, além de sugerir direções para trabalhos futuros.

Capítulo II Contextualização

A análise de séries temporais permite desvendar padrões e comportamentos em dados coletados ao longo do tempo. Esses dados, organizados em sequências cronológicas, possibilitam *insights* valiosos para uma variedade de campos, desde economia e saúde até previsão do tempo. Por meio da análise dessas séries, é possível identificar tendências, eventos importantes e relações entre variáveis [Esling and Agon, 2012; Shumway and Stoffer, 2017].

No entanto, a análise eficaz de séries temporais requer não apenas o entendimento dos dados em si, mas também a aplicação de técnicas adequadas de pré-processamento. Este capítulo aborda os principais conceitos relacionados à RT, desde as técnicas fundamentais para sua execução, como diferenciação, uso de janelas deslizantes e decomposição de séries temporais, até características importantes dos dados, como rugosidade e heteroscedasticidade. Ao longo deste capítulo, também é feita uma breve descrição sobre a detecção de anomalias em séries temporais e são apresentadas as métricas empregadas na avaliação experimental deste trabalho.

II.1 Séries temporais

Dados de séries temporais são sequências de medições ao longo do tempo que normalmente seguem um determinado comportamento [Esling and Agon, 2012]. Formalmente, uma série temporal Y é uma sequência de observações em intervalos de tempo, ordenadas cronologicamente, do tipo $\langle y_1, y_2, y_3, \ldots, y_n \rangle$, sendo n o número de observações. Uma observação específica dessa série é referenciada como y_t, sendo indexada no tempo por t ($t \in \{1, \ldots, n\}$). Para séries temporais amostradas em intervalos de tempo regulares, o tempo decorrido constante entre duas observações consecutivas é denominado intervalo de amostragem da série. O intervalo de amostragem pode variar desde frações de segundo, como, por exemplo, no caso de monitoramento de frequência cardíaca, até a frequência mensal, anual ou mais espaçada no tempo, como, por exemplo, em observações climatológicas.

A análise de séries temporais é utilizada para compreender e prever dados temporais em vários domínios. É valiosa para estudos como a identificação de tendências e sazonalidades dos dados, a detecção de eventos importantes, a previsão de comportamento futuro e o estudo de relações entre variáveis no tempo, entre outros [Esling and Agon, 2012]. Essas análises podem ser aplicadas em

diversas áreas, como economia (*e.g.*, avaliação de riscos financeiros), saúde e medicina (por exemplo, identificação de surtos em doenças epidemiológicas), comércio (*e.g.*, análise de tendência de vendas), tempo e clima (*e.g.*, previsão de desastres naturais), entre outras. Em resumo, a análise de séries temporais é uma ferramenta que gera descobertas a partir de dados observados ao longo do tempo [Shumway and Stoffer, 2017].

Para a análise das séries temporais, assim como para a análise de dados de forma geral, um adequado pré-processamento é uma atividade importante. O pré-processamento abrange grande parte do processo de mineração de dados. Inclui limpeza de dados, seleção de atributos, amostragem, normalização, transformação de dados e outras técnicas, buscando garantir a qualidade dos dados antes de serem submetidos a qualquer método de análise [Garcia et al., 2014; Han et al., 2022].

A transformação de dados envolve os processos que criam novas séries ou variáveis a partir da série original. O objetivo é capturar informações ou adaptá-las às condições necessárias à análise. É importante, nesse processo, que não se percam as características geradoras da série, sob pena de distorcer ou corromper os dados e levar a respostas enviesadas ou incoerentes. Exemplos de transformação de dados de séries temporais são estabilização de variância, redução de tendência, diferenciação e decomposição [Gujarati, 2021; Shumway and Stoffer, 2017].

II.1.1 Normalização

A normalização de dados é um método de transformação empregado na análise de dados para ajustar os valores de características numéricas a uma escala comum, preservando sua estrutura temporal. Na manipulação de dados de diferentes origens, essa técnica assegura que todas as características contribuam igualmente para a análise, prevenindo que aquelas com escalas maiores dominem as outras.

Existem várias técnicas para normalizar dados. A Equação II.1 mostra a normalização Z-Score, na qual $Z(Y_t)$ é a série temporal padronizada, Y_t é a série original, μ é a média e σ é o desvio padrão das observações da série. Essa normalização garante que os dados transformados tenham média igual a 0 e desvio padrão igual a 1.

$$Z(Y_t) = \frac{Y_t - \mu}{\sigma} \tag{II.1}$$

II.1.2 Diferenciação

A diferenciação envolve calcular a diferença entre observações consecutivas em uma série temporal. É uma técnica comumente usada na análise de dados de séries temporais para torná-los estacionários e facilitar a modelagem e a análise por meio de métodos estatísticos [Cryer and Chan, 2008; Shumway and Stoffer, 2017]. Na diferenciação de primeira ordem, subtrai-se cada observação

da observação anterior. Matematicamente, a diferenciação de primeira ordem é representada na Equação II.2, em que y_t é uma observação no tempo t. A série resultante ∇y_t possui uma observação a menos que a série original, pois o primeiro termo não tem um anterior para subtrair.

$$\nabla y_t = y_t - y_{t-1} \tag{II.2}$$

II.1.3 Janelas deslizantes

A análise de séries temporais com janelas deslizantes é uma técnica utilizada para examinar e prever comportamentos dinâmicos de dados ao longo do tempo [Lampert et al., 2008]. Essa abordagem envolve a criação de subconjuntos de dados em uma janela que "desliza" pela série temporal. Cada posição da janela representa um subconjunto dos dados, permitindo a análise em intervalos menores da série para avaliar propriedades locais [Ogasawara et al., 2010].

Dois parâmetros são importantes: o tamanho e o passo da janela. O tamanho da janela define a quantidade de dados que compõem cada janela, ou seja, o tamanho do subconjunto. O passo da janela corresponde à quantidade de pontos que a janela se desloca até o próximo subconjunto.

Formalmente, considere uma série temporal $\langle y_1, y_2, y_3, \ldots, y_n \rangle$ indexada no tempo por t, em que n é o número de observações. Considere também f a função à qual o subconjunto será submetido (*e.g.*, média). À medida que a janela de tamanho w se move ao longo da série temporal, a função f é aplicada aos dados dentro de cada janela, gerando uma sequência de saída denotada por $\langle f_1, f_2, f_3, \ldots, f_{n-w+1} \rangle$. Assim, f_i representa o resultado da aplicação de f à i-ésima janela, ou seja, $f_i = f(y_i, y_{i+1}, \ldots, y_{i+w-1})$, em que $i = 1, 2, \ldots, n - w + 1$.

II.1.4 Heteroscedasticidade

No contexto de séries temporais, a heteroscedasticidade refere-se à condição em que a variância dos erros ou resíduos não é constante ao longo do tempo, contrastando com as séries homoscedásticas, nas quais a variância dos dados permanece homogênea. Esta propriedade é fundamental para a análise de séries temporais, mas exige uma distinção clara entre suas formas incondicional e condicional.

A heteroscedasticidade incondicional ocorre quando a variância de longo prazo ou global de um processo estocástico flutua sem um padrão definido ou exibe tendências claras, implicando que a série não é estacionária de segunda ordem. Por exemplo, uma alteração drástica na qualidade da coleta de dados pode levar a uma diminuição significativa na variação do erro ao longo do tempo, resultando em heteroscedasticidade incondicional.

Em contrapartida, a heteroscedasticidade condicional é caracterizada por uma variância que muda no tempo de forma previsível, dependendo das informações passadas da série. Ela se mani-

festa quando a variância de uma observação no tempo, considerando a informação disponível até o momento anterior, não é constante. Este fenômeno é frequentemente observado por agrupamentos de variabilidade, onde períodos de alta variabilidade tendem a ser seguidos por outros de alta variabilidade, e o mesmo ocorre para períodos de baixa variabilidade. É importante notar que um processo pode apresentar heteroscedasticidade condicional e ainda assim ser estacionário de segunda ordem, desde que sua variância incondicional (de longo prazo) seja constante. A prevalência da heteroscedasticidade, especialmente em sua forma condicional, é notória em áreas como finanças, onde os retornos de ativos exibem essa dinâmica de volatilidade. Modelar e prever essa volatilidade é essencial para análise de risco, estudo de mercado, investimentos e outras aplicações [Wooldridge, 2013].

A heteroscedasticidade é frequentemente discutida em relação à suposição de análises paramétricas, como a regressão linear, pois invalida testes estatísticos de significância que pressupõem que todos os erros de modelagem tenham a mesma variância. Isso pode comprometer a validade e a confiabilidade em estudos de séries temporais, incluindo, além de modelos de regressão, métodos de mínimos quadrados e testes estatísticos. Assumir homoscedasticidade quando, na verdade, há heteroscedasticidade resulta em estimativas pontuais imparciais, porém ineficientes, e em estimativas tendenciosas de erros padrão, podendo levar à superestimação da qualidade do ajuste realizado por vários coeficientes. Por isso, detectar e abordar a heteroscedasticidade é essencial para obter resultados precisos e significativos [Wooldridge, 2013].

Além da análise visual, existem testes estatísticos específicos para verificar a presença de heteroscedasticidade em séries temporais, como o teste de White [White, 1980], o teste de Bartlett [Bartlett, 1937], o teste Breusch–Pagan [Breusch and Pagan, 1979], o teste de Goldfeld–Quandt [Goldfeld and Quandt, 1965] e o teste de Brown–Forsythe [Brown and Forsythe, 1974]. Existem muitos outros testes para heteroscedasticidade. Esses testes fazem diferentes suposições sobre a natureza da heteroscedasticidade e sobre as características das séries, como a distribuição dos termos de erro, entre outros [Wooldridge, 2013]. Esses testes analisam a variação dos resíduos em vários subgrupos dos dados ou em diferentes funções das variáveis independentes e rejeitam a hipótese nula de homoscedasticidade se detectarem diferença significativa.

Caso seja comprovada a heteroscedasticidade na série de interesse, existem formas de mitigar essa característica a partir de transformações. As transformações conseguem estabilizar a variação (transformar heteroscedasticidade em homoscedasticidade). Podem ser logarítmicas, de raiz quadrada, Box–Cox [Box and Cox, 1964], etc. Embora as transformações possam ser eficazes na abordagem da heteroscedasticidade, nem todas as séries podem ser transformadas adequadamente (sem perder a confiabilidade da análise). Além disso, as transformações devem ser utilizadas criteriosamente, pois podem tornar as unidades imprecisas e dificultar a interpretação dos resultados,

além de comumente exigirem amplo conhecimento na área e muitas manipulações nos dados da série temporal de interesse [Salles et al., 2019].

Como alternativa às transformações, existem modelos que consideram explicitamente a heteroscedasticidade, como os modelos *Autoregressive Conditional Heteroskedasticity* (ARCH) [Engle, 1982]. Esses modelos não tentam corrigir a heteroscedasticidade, mas implementam recursos adaptados a essa característica. São amplamente utilizados para modelar séries temporais, utilizando-se de equações de variância condicional capazes de descrever como a volatilidade evolui ao longo do tempo e de parâmetros que monitoram a persistência da volatilidade. A comunidade científica desenvolveu diversas extensões para o modelo clássico ARCH, como o modelo *Generalized Autoegressive Conditional Heteroskedasticity* (GARCH) [Bollerslev, 1986], o modelo *Nonlinear Asymmetric Generalized Autoegressive Conditional Heteroskedasticity* (NAGARCH) [ENGLE and NG, 1993], o modelo *Exponencial Generalized Autoegressive Conditional Heteroskedasticity* (EGARCH) [Nelson, 1991], o modelo *Glosten-Jagannathan-Runkle Generalized Autoegressive Conditional Heteroskedasticity* (GJR-GARCH) [GLOSTEN et al., 1993], o modelo *Quadratic Generalized Autoegressive Conditional Heteroskedasticity* (QGARCH) [Sentana, 1995], entre outros. Esses métodos incrementais incorporam recursos adicionais para capturar diferentes aspectos da dinâmica da volatilidade.

II.1.5 Rugosidade

No estudo de séries temporais, a caracterização da variabilidade e a identificação de irregularidades nos dados são essenciais para a compreensão de seus padrões. Dentro desse escopo, a rugosidade é uma métrica que quantifica o grau de variabilidade ou irregularidade nos dados de uma série temporal [Sarkar, 2006]. Nesse contexto, ela representa uma medida da irregularidade local da série. É uma adaptação do conceito clássico de rugosidade média, amplamente utilizado na engenharia de superfícies para descrever a variabilidade de perfis [Souza et al., 2024; Pour, 2018]. Em superfícies físicas, uma maior rugosidade corresponde a uma maior variação de altura ao longo do perfil, indicando maior frequência e amplitude nas oscilações locais. Analogamente, em séries temporais, uma série mais rugosa exibe oscilações mais rápidas e abruptas, frequentemente associadas a componentes de alta frequência [Song et al., 2014]. Tais irregularidades são de particular interesse na detecção de anomalias, pois eventos anômalos muitas vezes se manifestam como desvios súbitos e de curta duração que aumentam significativamente a irregularidade local da série [Chandola et al., 2009].

Entender a rugosidade de uma série temporal é importante para escolher técnicas adequadas e compreender a dinâmica subjacente dos dados. Por exemplo, séries temporais mais suaves podem ser melhor tratadas por modelos estatísticos convencionais, enquanto séries mais rugosas podem demandar métodos mais avançados, capazes de capturar padrões e flutuações complexas [Box et al.,

2015]. No contexto da detecção de anomalias, a análise da rugosidade se mostra valiosa para identificar componentes da série que capturam as variações mais abruptas e imprevisíveis, características de muitos eventos anômalos. Componentes com maior rugosidade tendem a ser mais sensíveis a essas perturbações localizadas, tornando-os alvos prioritários para algoritmos de detecção de anomalias [Zamanzadeh Darban et al., 2024].

A rugosidade nos dados de séries temporais pode ser quantificada usando várias medidas estatísticas ou técnicas matemáticas. A rugosidade γ é calculada de acordo com a Equação II.3, sendo definida como a média das variações quadráticas locais das primeiras diferenças da série, previamente normalizadas pelo seu desvio padrão.

$$\gamma = \frac{1}{n-2} \sum_{t=2}^{n-1} \left(\frac{\nabla y_{t+1} - \nabla y_t}{\sigma_\nabla} \right)^2 \cdot \frac{1}{4} \tag{II.3}$$

Nesta equação, ∇y_t denota a primeira diferença da série temporal, e σ_∇ é o desvio padrão das primeiras diferenças da série. O somatório é executado de $t = 2$ até $n - 1$, garantindo que todos os termos sejam bem definidos dentro do intervalo da série. A expressão dentro do somatório quantifica o quanto a inclinação da série muda de um passo de tempo para o seguinte, o que está diretamente relacionado à sua curvatura local. O fator $\frac{1}{4}$ deriva da formulação clássica da curvatura discreta em séries com espaçamento unitário, assegurando consistência com as medidas geométricas de variação de inclinação. Dessa forma, γ serve como um indicador da rugosidade de segunda ordem de uma série temporal.

A escolha desta formulação para a rugosidade é estratégica, especialmente em análises que envolvem séries temporais heteroscedásticas. A normalização pelo desvio padrão das primeiras diferenças (σ_∇) confere à métrica independência de escala, assegurando que o valor de rugosidade não seja indevidamente influenciado por variações na magnitude ou variância da série ao longo do tempo. Além disso, o uso das segundas diferenças (implícito em $\nabla y_{t+1} - \nabla y_t$) torna a métrica robusta a tendências globais, isolando as irregularidades locais sem confundir a rugosidade com variações de nível ou inclinação de longo prazo [Box et al., 2015]. Essa característica é fundamental para a comparação equitativa da rugosidade entre diferentes segmentos de uma série ou entre diferentes componentes.

II.1.6 Curvatura máxima

A curvatura máxima se refere ao ponto de uma função em que a curvatura é maior. Ela mede quão rapidamente uma curva muda de direção em um determinado ponto. A curvatura máxima tem aplicações em vários campos, incluindo geometria, computação gráfica, engenharia e medicina. Em matemática, estudar a curvatura máxima pode auxiliar na compreensão das propriedades e do

comportamento de funções.

No contexto da análise de sequências, a curvatura máxima pode ser aplicada para identificar pontos cruciais em que o comportamento da sequência se altera significativamente. Para identificar esse ponto de transição, a curvatura de uma sequência y, com elementos y_j para $1 \leq j \leq k$, é estimada utilizando um *spline* suavizado (Equação II.4). Essa técnica permite uma computação estável da segunda derivada de y. O índice d é então definido como o ponto de curvatura máxima, ou seja, o valor de j que maximiza y''_j:

$$d = \underset{1 \leq j \leq k}{\arg\max} \left(y''_j \right). \tag{II.4}$$

O índice d representa um ponto de transição na sequência y, marcando uma mudança significativa em seu comportamento intrínseco. A identificação desse ponto permite delimitar uma porção da sequência y para análise ou processamento direcionado.

II.2 Decomposição de séries temporais

A decomposição de séries temporais é um recurso utilizado para analisar a série em vários componentes, cada um representando um padrão subjacente dos dados. Por definição, a partir dos componentes é possível reproduzir a série original, seja pela multiplicação (modelo de decomposição multiplicativo), seja pela adição (modelo de decomposição aditivo) desses componentes [Box et al., 2015]. Os componentes são comumente organizados em três tipos principais: tendência, sazonalidade e resíduo [Hyndman and Athanasopoulos, 2018]. Porém, existem outros modos de decomposição.

A decomposição de modo empírico EMD é um método de decomposição tempo-frequência aditivo, introduzido por Huang et al. [1998]. Ela decompõe as séries nas chamadas funções do modo intrínseco (IMF), que representam diferentes modos oscilatórios da série temporal. As IMF obedecem a dois principais requisitos: (1) números iguais de cruzamentos de zero e de extremos; e (2) envelopes simétricos. Isso faz com que elas sejam mais estáveis, mais bem comportadas e, consequentemente, mais fáceis de trabalhar. A EMD é particularmente útil para séries não estacionárias, nas quais os componentes de frequência mudam com o tempo [Zeiler et al., 2010]. A Equação II.5 mostra a série Y indexada no tempo por t, que é decomposta em k IMFs (ι) mais o resíduo ρ.

$$Y_t = \sum_{i=1}^{k} \iota_t^i + \rho_t \tag{II.5}$$

Para extrair as IMFs, é utilizado um processo de peneiração iterativo. Para cada IMF, primeiro identificam-se os extremos locais (máximos e mínimos), criam-se envelopes (superior e inferior) a partir da interpolação dos extremos, calcula-se a média desses envelopes e subtrai-se essa média da série original. A série resultante é avaliada para verificar se atinge os requisitos de saída (ou seja, se

possui o mesmo número de cruzamentos no zero e nos extremos e se possui envelopes simétricos). Se isso não ocorrer, o processo se repete usando a nova série como entrada. Se os requisitos forem atingidos, a IMF é gerada, e o que sobra da série original passa por toda a iteração novamente, até que todas as IMF sejam geradas. O resíduo é o componente que representa o que sobra da série original, a partir do qual não é possível gerar mais nenhuma IMF [Huang et al., 1998].

A EMD extrai recursivamente diferentes resoluções dos próprios dados, sem o uso de funções ou filtros fixos. Isso é uma vantagem em relação a outras técnicas de decomposição, que podem exigir muitos parâmetros de entrada. Usando a EMD, é possível eliminar componentes específicos, como o resíduo, e reconstruir a série. Também é possível extrair componentes específicos para análise direcionada, por exemplo, usar as IMF de maior frequência para detecção de anomalias ou as de menor frequência e/ou o resíduo para determinação de tendência.

O algoritmo EMD apresenta um problema de mistura de modos de frequência. Ou seja, sinais com escalas semelhantes, que deveriam estar na mesma IMF, acabam sendo alocados em outra IMF. Isso pode causar falseamento (*aliasing*) na distribuição tempo-frequência e interferir no significado estatístico das IMFs. Para minimizar esse problema, Wu and Huang [2009] propuseram um aprimoramento do EMD: *Ensemble Empirical Mode Decomposition* (EEMD). O EEMD é um método de análise assistida por ruído, que adiciona ruído branco aos dados originais para forçar o conjunto a esgotar todas as soluções possíveis no processo de peneiração. Dessa forma, as escalas de frequência da série são separadas naturalmente.

Apesar dos avanços trazidos pelo EEMD na mitigação da mistura de modos, um desafio remanescente estava associado ao seu alto custo computacional e à persistência de ruído residual após o processo de decomposição. Para otimizar ainda mais essa abordagem, Yeh et al. [2010] propuseram um aprimoramento significativo, introduzindo o *Complete Ensemble Empirical Mode Decomposition* (CEEMD). O CEEMD visa mitigar o problema do ruído residual e o alto custo computacional do EEMD por meio do uso de pares de ruído branco (positivo e negativo) adicionados à série original. Esse método gera menos testes e ajustes iterativos, resultando em um processo mais eficiente, sem comprometer a capacidade de decomposição e a separação das escalas de frequência da série. A Figura II.1 mostra um exemplo da decomposição usando o CEEMD para uma série temporal sintética.

II.3 Detecção de eventos

Na análise de séries temporais, a detecção de eventos é um dos objetivos centrais, sendo estes compreendidos como ocorrências significativas ou desvios do comportamento esperado das observações. Tais eventos podem se manifestar predominantemente como anomalias, pontos de mudança ou *motifs* [Ogasawara et al., 2025b]. Um *motif* representa um padrão ou subsequência que se repete

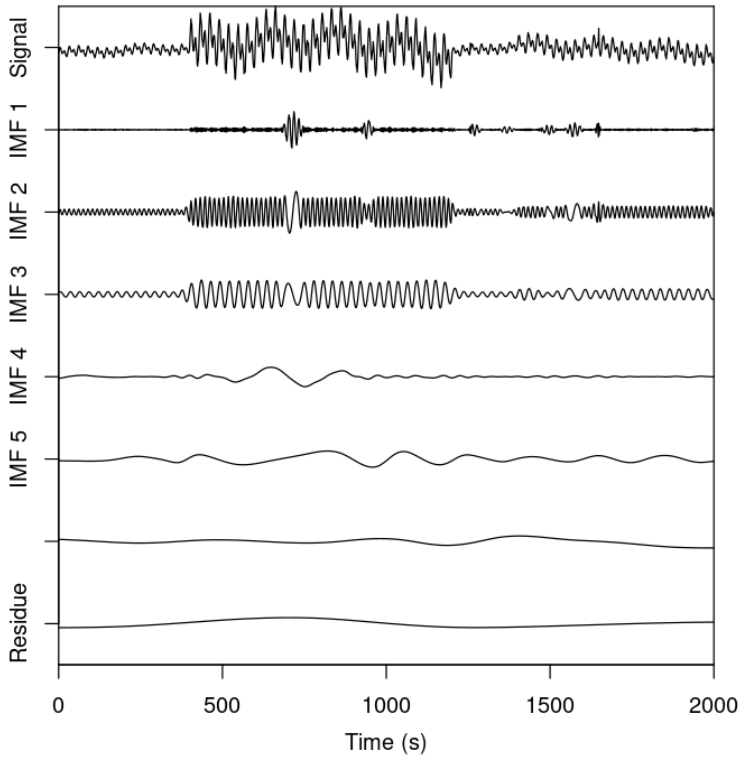


Figura II.1: Decomposição utilizando CEEMD de série temporal sintética, sendo a primeira série (superior) a série original, seguida pelas seis IMF geradas e, por último (inferior), a série dos resíduos.

frequentemente na série, revelando comportamentos recorrentes intrínsecos ao sistema que gera a série. Em contraste, tanto as anomalias quanto os pontos de mudança sinalizam alguma forma de desvio ou transição no comportamento da série temporal em relação a um padrão estabelecido ou esperado.

Um ponto de mudança é uma instância no tempo em que as propriedades estatísticas subjacentes de uma série temporal, como média, variância ou tendência, sofrem uma alteração abrupta e persistente, indicando uma transição para um novo regime ou estado. Por sua vez, a anomalia é caracterizada como uma observação ou um conjunto de observações que se desvia significativamente do padrão normal ou do comportamento majoritário da série temporal [Chandola et al., 2009].

Existem vários métodos e algoritmos empregados para detectar eventos em dados de séries temporais, como métodos estatísticos, abordagens de aprendizado de máquina e técnicas de processamento de sinais [Esling and Agon, 2012; Box et al., 2015; Shumway and Stoffer, 2017]. Cada uma dessas técnicas tem seus pontos fortes e limitações, e a escolha do método de detecção apropriado não é trivial. Ela depende de fatores como as características dos dados, a natureza dos eventos, os recursos computacionais disponíveis e os objetivos da pesquisa [Ogasawara et al., 2025b; Zou et al., 2023; Wenig et al., 2024].

A seleção de hiperparâmetros apropriados nos métodos de detecção representa um desafio significativo. Muitos algoritmos de detecção exigem a especificação de determinados parâmetros, es-

pecialmente os modelos mais elaborados, que envolvem aprendizado de máquina com parâmetros de ajuste. Uma pequena variação nesses parâmetros pode afetar drasticamente o desempenho de um método. Faz-se necessário, também, procurar evitar sobreajuste ou subajuste e encontrar o equilíbrio certo entre sensibilidade (a capacidade de detectar eventos verdadeiros) e especificidade (a capacidade de evitar alarmes falsos). Ainda, a seleção de parâmetros pode exigir conhecimento de domínio ou experiência na área de aplicação específica. Como os parâmetros podem afetar significativamente a precisão, a eficiência e a robustez do processo de detecção, quando o método exige a determinação de muitos e complexos parâmetros, sua utilização torna-se mais difícil e seus resultados mais instáveis [Chandola et al., 2009; Esling and Agon, 2012; Ogasawara et al., 2025b].

II.3.1 Anomalias

Anomalias em séries temporais são eventos que correspondem a padrões inesperados ou anormais, que parecem não ter sido gerados pelo mesmo processo que gerou o restante dos dados da série temporal [Chandola et al., 2009]. Podem estar presentes por diversos fatores, desde problemas na coleta dos dados até eventos incomuns importantes, como fraude em cartão de crédito.

Existem várias formas de classificar anomalias. Uma classificação é quanto ao tipo, que pode ser pontual ou de sequência. Anomalias pontuais são observações únicas, e anomalias de sequência correspondem a uma sequência de observações que se destacam significativamente do restante dos dados [Boniol and Palpanas, 2020]. Outra classificação é quanto à circunstância, podendo ser contextual ou descontextualizada [Ogasawara et al., 2025b; Zou et al., 2023]. Uma anomalia descontextualizada ocorre quando o valor é totalmente incoerente no contexto do dado observado, como, por exemplo, uma temperatura de 100 °C em uma série de monitoração da temperatura de um ser humano. Já a anomalia contextual representa um valor discrepante somente no contexto daquele intervalo da série em que se encontra, mesmo que esse seja um valor considerado razoável para ser atribuído às observações da série. Um exemplo de anomalia contextual é um termômetro apresentar leitura de 30 °C à noite durante a primavera em uma cidade com clima temperado, onde são esperadas temperaturas amenas durante a noite, geralmente entre 10 °C e 15 °C. Ou seja, embora 30 °C não seja uma temperatura impossível (em outro local ou em outro período), ela é anômala para o contexto.

Compreender os diferentes tipos de anomalias é importante para o desenvolvimento de métodos e estratégias eficazes de detecção que sejam adaptados às características específicas dos dados das séries temporais e à natureza das anomalias presentes.

II.4 Métricas de avaliação

Dada a diversidade de métodos de detecção disponíveis na literatura, é extremamente importante avaliar sua capacidade em detectar corretamente todos os eventos e, também, em não apontar eventos erroneamente. Para isso, são utilizadas métricas de avaliação. As métricas fornecem medidas quantitativas do desempenho de um método de detecção, permitindo avaliar e comparar diferentes métodos e suas configurações [Sullivan and Feinn, 2012].

A matriz de confusão é uma tabela que sumariza o desempenho de um modelo de classificação ou detecção, confrontando as predições do modelo com os rótulos reais (ou gabarito). As métricas clássicas de detecção se baseiam diretamente em seus elementos. Tais elementos são: *falso positivo* (FP), *falso negativo* (FN), *verdadeiro positivo* (VP) e *verdadeiro negativo* (VN). Eles resumem os resultados do emprego do método, comparando os eventos encontrados com os rótulos reais. Dessa forma, só é possível aplicar essas métricas quando existe um gabarito, indicando quais elementos da série temporal são, de fato, eventos. Os "falsos" são erros na detecção, enquanto os "verdadeiros" são acertos. Assim, quanto maiores os VP e VN, melhor é o desempenho do método em detectar corretamente os eventos. Os FP são conhecidos como alarmes falsos por determinarem evento onde não há, e os FN são os eventos perdidos. É interessante visualizar a contagem desses elementos em porcentagens, já que o número de exemplos em cada classe pode variar.

Como mencionado anteriormente, os valores de FP, FN, VP e VN servem como base para o cálculo de várias métricas de desempenho, como precisão, revocação e F1-score. Como pode ser visto na Equação II.6, a precisão mede a proporção de resultados VP entre todos os resultados positivos previstos pelo modelo (VP e FP). A precisão é particularmente importante quando o custo dos FP é significativo. Varia em uma escala de 0 a 1, sendo 1 quando acerta todos os eventos. Isso não quer dizer, entretanto, que o método tenha detectado todos os eventos rotulados (pois a revocação pode ser baixa).

$$precisão = \frac{VP}{VP + FP} \tag{II.6}$$

Já a revocação é a métrica que mede a frequência com que um método detecta corretamente os verdadeiros positivos de todos os eventos rotulados na série temporal. A Equação II.7 mostra como a revocação é calculada a partir da divisão do número de VP pelo número de instâncias positivas (VP e FN). É importante quando o custo dos FN é alto. Varia em uma escala de 0 a 1, sendo 1 quando o modelo consegue detectar todos os eventos rotulados. Por exemplo, supondo que um método retorne que todos os valores de uma série temporal são eventos, a revocação é 1, e isso não significa que o desempenho do método foi bom. Nesse caso, a precisão é muito baixa.

$$revocação = \frac{VP}{VP + FN} \tag{II.7}$$

Dependendo do objetivo da pesquisa, prioriza-se aumentar uma métrica ou outra. Por exemplo, caso se trate de eventos que configuram grande risco, é desejável que todos os eventos verdadeiros sejam detectados, mesmo que falsos positivos sejam mais comuns. Nesse caso, a métrica que deve estar alta é a revocação. Por outro lado, ao se tratar de uma aplicação na qual os FP são considerados mais prejudiciais que os FN, a precisão é que deve estar alta. Um exemplo disso é quando é necessária uma certa garantia de que o evento classificado é, de fato, um evento, como na decisão por aplicar em um investimento. Nesse caso, garantir o acerto é muito mais importante, mesmo que outras oportunidades (VP) sejam perdidas.

A métrica F1 consegue avaliar conjuntamente tanto a precisão quanto a revocação. Como pode ser visto na Equação II.8, F1 é a média harmônica entre as outras duas métricas. Então, para que o F1 seja alto, tanto a precisão quanto a revocação devem ser altas. Essa métrica também varia de 0 a 1, podendo ser indeterminada caso precisão e revocação sejam iguais a zero. Valores de F1 próximos ou iguais a 1 significam que o método obteve um bom desempenho tanto em acertar suas predições (precisão alta) quanto em detectar o maior número de eventos rotulados (revocação alta. Pelas suas características, essa métrica tende a ser um resumo da qualidade do modelo. Por outro lado, pode ser difícil interpretá-la sozinha e, por isso, é aconselhável observar conjuntamente F1, precisão e revocação.

$$F1 = \frac{2 \times precisão \times revocação}{precisão + revocação} = \frac{2VP}{2VP + FP + FN} \tag{II.8}$$

II.4.1 Testes Estatísticos e Tamanho do Efeito

A validação de hipóteses em pesquisas científicas frequentemente emprega testes estatísticos para inferir características populacionais a partir de dados amostrais [Fagerland, 2012]. Muitos desses testes são paramétricos, assumindo que os dados seguem distribuição normal e as variâncias dos grupos são homogêneas. Contudo, em diversas áreas do conhecimento essas premissas podem não ser atendidas e, nesses cenários, os testes não paramétricos são mais adequados por não dependerem de suposições sobre a forma da distribuição ou da homogeneidade da variância [Cox, 1982]. Essa flexibilidade os torna particularmente adequados para lidar com dados que exibem assimetria ou naturezas ordinais, garantindo a validade das inferências mesmo sob condições menos ideais.

A avaliação da significância estatística é comumente realizada por meio do *p-valor*, que quantifica a probabilidade de observar resultados tão extremos quanto (ou mais extremos do que) os obtidos, sob a premissa de que a hipótese nula é verdadeira. Um baixo p-valor (*e.g.*, $p < 0{,}05$) sugere que

os dados fornecem evidências para rejeitar a hipótese nula, indicando que a diferença ou relação observada provavelmente não ocorreu por mero acaso amostral [Tomczak and Tomczak, 2014]. A definição da hipótese alternativa, que representa o que o pesquisador busca investigar, pode ser bilateral (ou bicaudal), quando se interessa por qualquer diferença entre os grupos, ou unilateral (ou unicaudal), quando se busca uma diferença em uma direção específica (*e.g.*, um grupo é estritamente maior que outro) [Cox et al., 1977]. Para determinar a existência de uma diferença sem especificar a direção, um teste bicaudal é apropriado. No entanto, é crucial reconhecer as limitações do p-valor: ele não informa sobre a magnitude ou importância prática do efeito observado [Sullivan and Feinn, 2012].

Para complementar o p-valor e fornecer uma compreensão mais completa dos resultados, o tamanho do efeito torna-se indispensável. O tamanho do efeito é uma medida padronizada da magnitude de uma diferença ou da força de uma associação [Hojat and Xu, 2004; Sullivan and Feinn, 2012]. Ele responde à questão de quanto um efeito é substancial, distinguindo a significância estatística da significância prática [Hojat and Xu, 2004]. Em contextos como comparação de desempenhos de métodos, o tamanho de efeito pode indicar não apenas que há uma diferença de desempenhos, mas que essa distinção possui uma magnitude relevante. A interpretação do tamanho do efeito permite, assim, avaliar a importância real dos achados, facilitando comparações entre diferentes estudos e metodologias [Tomczak and Tomczak, 2014].

Para a comparação de distribuições entre duas amostras, como na avaliação do desempenho de métodos de detecção de eventos ou anomalias, o teste de Wilcoxon é frequentemente empregado [Nirmalkar, 2025; Bauer, 1972]. Esse é um teste não paramétrico robusto que avalia se a diferença entre as medianas de duas amostras é estatisticamente significativa. Para quantificar a magnitude dessas diferenças, utiliza-se o cálculo do tamanho do efeito associado ao teste de Wilcoxon, expresso pelo r de Rosenthal [Rosenthal and Rubin, 2003]. Uma característica fundamental desse cálculo é que ele se baseia nos ranques dos dados originais, e não em seus valores brutos, proporcionando uma métrica que reflete a sobreposição das distribuições. A interpretação do tamanho do efeito segue a convenção de Cohen [Rosenthal, 1994; Sullivan and Feinn, 2012; Cohen, 2013]: valores até 0,1 indicam um efeito pequeno (de importância prática negligenciável), 0,3 um efeito médio (moderada importância prática) e 0,5 ou mais um efeito grande (crucial importância prática). Essa abordagem conjunta permite não apenas verificar a existência de uma distinção estatística, mas também compreender a relevância prática dessa distinção, um aspecto fundamental para a avaliação precisa e aplicável em validações metodológicas.

Capítulo III Trabalhos Relacionados

A seleção dos trabalhos analisados nesta seção foi realizada a partir de uma busca sistemática na plataforma *Elicit*, conduzida em 18/07/2025. A *string* de pesquisa utilizada foi formulada com base nos termos em inglês das três dimensões que nortearam o estudo e que guiam a formulação do problema apresentada na introdução desta Tese: *heteroscedasticidade*, *detecção de anomalias* e *transformação*. Os resultados retornados pelo *Elicit* foram então filtrados conforme essas mesmas dimensões. Essa organização permitiu abranger, de forma estruturada, os trabalhos mais alinhados aos eixos centrais desta pesquisa. Além disso, aplicou-se a técnica de *snowballing* a partir dos artigos inicialmente selecionados, de modo a complementar a busca e garantir a inclusão de estudos relevantes não capturados na etapa inicial. Os próximos parágrafos apresentam os trabalhos mais relacionados ao escopo desta pesquisa, compondo uma revisão estruturada da literatura. A detecção de anomalias em séries temporais é uma tarefa intrinsecamente desafiadora, e essa complexidade é significativamente ampliada pela presença de heteroscedasticidade, isto é, a variabilidade da dispersão dos dados ao longo do tempo. Em ambientes heteroscedásticos, a própria definição de normalidade e anomalia torna-se fluida. Um valor que seria considerado atípico em um período de baixa variância pode ser perfeitamente normal em um contexto de alta dispersão, tornando a linha de base dinâmica e difícil de estabelecer. Consequentemente, a aplicação de limiares fixos ou métricas de detecção estáticas torna-se inadequada, levando a uma elevada taxa de falsos positivos (alertas desnecessários em momentos de alta variância) ou falsos negativos (perda de anomalias genuínas em períodos de baixa variância, em que um desvio menor já seria significativo). Embora limiares adaptativos sejam uma alternativa necessária, sua calibração e generalização para diferentes séries e cenários de heteroscedasticidade representam um desafio metodológico considerável.

Ademais, muitos modelos clássicos e até mesmo algumas abordagens modernas de detecção de anomalias operam sob a premissa de homoscedasticidade ou estacionariedade, o que significa que a violação dessas suposições pode levar à subestimação ou superestimação da incerteza e, consequentemente, a erros de modelagem e detecção. Um desafio particular reside na preservação da estrutura temporal e contextual das anomalias. Transformações genéricas aplicadas para estabilizar a variância, por exemplo, correm o risco de distorcer ou até mesmo eliminar informações cruciais que definem uma anomalia como contextual, isto é, um evento que só é anômalo em um determinado momento ou padrão. A capacidade de distinguir e tratar anomalias pontuais, contextuais e coletivas,

sem comprometer a essência do comportamento temporal da série, é, portanto, um ponto crítico na formulação de detectores robustos em cenários heteroscedásticos.

A preocupação com a influência da heteroscedasticidade nos resultados da detecção em séries temporais é amplamente debatida na literatura [McArdle and Anderson, 2004; Jiang et al., 2009; Hassan et al., 2012; Loperfido, 2018; Jiang et al., 2023; Mari and Baldassari, 2021], destacando que a variabilidade temporal pode distorcer limiares fixos e prejudicar a identificação de eventos locais. Nesse sentido, reforça-se a importância de incorporar mecanismos para detecção ou adaptação à variância variável [Antunes et al., 2017; Cavaliere, 2004; Salles et al., 2019], de forma a preservar a sensibilidade dos métodos frente a mudanças estruturais nas séries.

Um dos caminhos mais tradicionais para lidar com esse problema está nos modelos da classe *Generalized Autoregressive Conditional Heteroskedasticity* (GARCH) e suas extensões, amplamente utilizados para modelar a variância condicional [Francq and Zakoian, 2019; Bollerslev, 1986]. Esses modelos permitem capturar oscilações locais ao longo do tempo, tendo se consolidado em aplicações financeiras devido à sua capacidade de representar *clusters* de volatilidade. Trabalhos como os de Spokoiny [2009] também abordam a modelagem e previsão de séries temporais não estacionárias com foco na volatilidade, utilizando o conceito de homogeneidade local. No entanto, a eficácia dos modelos GARCH depende de forte parametrização, além de assumir estrutura específica para o comportamento da variância, o que pode limitar a aplicação em contextos distintos e em séries não estacionárias [Bollerslev, 1986; Charles and Darné, 2019; Horváth et al., 2006].

Alternativamente, alguns estudos enfocam diretamente o diagnóstico de variância não constante e propõem transformações que estabilizam a dispersão da série. Trabalhos como os de Ridley [1994], Ogasawara et al. [2010], Riani [2009] e Wang et al. [2021] apresentam estratégias diversas de ajuste da variabilidade, seja por normalizações locais, seja por testes robustos para identificar heterogeneidade. Embora relevantes para o tratamento da heteroscedasticidade, essas abordagens não são integradas a *pipelines* completos de detecção de anomalias e, em geral, não consideram de forma explícita a preservação da estrutura temporal associada a eventos contextuais.

Cabe destacar que, entre os trabalhos revisados, poucos exploram de forma sistemática técnicas de decomposição, como EMD, *wavelets* ou decomposições sazonais, aplicadas diretamente ao problema de detecção de anomalias em séries heteroscedásticas.

Nos últimos anos, emergiram também métodos mais complexos que visam à robustez frente à não estacionariedade e à heteroscedasticidade. Entre eles, destacam-se modelos baseados em redes adversariais generativas [Xu et al., 2022; Geiger et al., 2020], como o proposto por Zhang et al. [2023] para detecção de anomalias multivariadas em Sistemas de Controle Industrial, combinando arquiteturas do tipo *Transformer* e redes adversárias generativas. Outras abordagens incluem arquiteturas neurais [Qiu et al., 2021], variações de *autoencoder* [Zhang et al., 2025; Wang and Zhang,

2025; Kim et al., 2024; Kieu et al., 2022], abordagens com memória de longo prazo [Lin et al., 2020; Kim and Kim, 2023] e métodos baseados em grafos [Chen et al., 2022b]. Trabalhos mais recentes no campo do aprendizado de máquina e aprendizado profundo para detecção de anomalias em dados de alta dimensão e complexos incluem uma variedade de modelos, como redes de memória de longo e curto prazo, redes neurais convolucionais, redes neurais de grafos, *Transformers* e algoritmos de aprendizado de máquina mais tradicionais, como k-vizinhos mais próximos [Hong et al., 2012; Au Yeung et al., 2020; Cheong et al., 2021; Tuli et al., 2022; Chen et al., 2022a; Alsan et al., 2023; Dong et al., 2023]. Alguns estudos, como o de Ou et al. [2023], propuseram métodos que utilizam Redes Convolucionais Temporais para estimar a correlação de longo alcance de séries temporais, apoiando a detecção de anomalias. Apesar de apresentarem bons resultados em cenários desafiadores, esses detectores costumam demandar grande volume de dados para treino supervisionado, parametrização extensa e elevado custo computacional, fatores que dificultam sua adoção em ambientes operacionais ou com recursos limitados.

Além disso, métodos estatísticos têm sido explorados para a detecção de comportamentos atípicos em séries temporais. Por exemplo, Balestrassi et al. [2008] introduziram um método para detectar mudanças dinâmicas em séries temporais não lineares usando gráficos de controle multivariados. Já Wang et al. [2006] utilizaram uma estrutura baseada em similaridade, empregando o algoritmo *Hierarchical Agglomerative Clustering*, para identificar valores discrepantes em séries temporais de alta dimensão, especialmente em mercados financeiros.

Em síntese, a literatura revisada evidencia três caminhos principais: (i) a modelagem estatística da variância condicional por meio de modelos GARCH e derivados, eficazes mas dependentes de forte parametrização; (ii) transformações que buscam estabilizar a dispersão, úteis para normalização mas pouco exploradas em *pipelines* completos de detecção; e (iii) detectores modernos baseados em arquiteturas complexas, robustos mas de alto custo computacional e baixa interpretabilidade. Nenhuma dessas abordagens, contudo, integra de forma conjunta a normalização da heteroscedasticidade, a preservação da estrutura temporal de eventos contextuais e a possibilidade de decisão simples após a transformação. Nesse contexto, a contribuição desta Tese é propor uma Transformação Resiliente (RT) que preenche essa lacuna, permitindo tanto a utilização como etapa de pré-processamento quanto como método autônomo de detecção (RTAD). Além do avanço conceitual, destaca-se a disponibilização do código-fonte completo e implementado, aspecto pouco frequente na literatura, em que a ausência de repositórios abertos frequentemente limita a reprodutibilidade. A simplicidade operacional da RT/RTAD contrasta com a complexidade de métodos recentes, reforçando sua aplicabilidade em contextos reais.

Entre os diversos estudos revisados, alguns apresentam elevada similaridade estrutural e conceitual com a RT e seu uso para detecção de anomalias (RTAD). Os trabalhos são: REMD [Souza

et al., 2024], a abordagem de Audibert et al. [2021], ATAD [Dani et al., 2015], WANEH [Kanarachos et al., 2017] e o método de pré-processamento de Sun et al. [2021].

O *Refined Empirical Mode Decomposition* (REMD), proposto por Souza et al. [2024] (que compartilha autoria com este trabalho), é um método híbrido de detecção de anomalias que integra a Decomposição Empírica de Modos (EMD), utilizando especificamente a CEEMD para separar componentes de frequência, seguida de uma etapa de seleção de componentes baseada na rugosidade e na curvatura máxima. Após essa separação, um modelo ARIMA é aplicado especificamente ao componente de alta frequência para identificar anomalias nos resíduos. Esse trabalho representa uma aplicação prévia de conceitos que mais tarde fundamentariam a RT, validando a ideia de decomposição e seleção de componentes para aprimorar a detecção de anomalias. O bom desempenho dessa abordagem, evidenciado em diferentes experimentos, corrobora a utilidade prática da combinação entre decomposição e realce de componentes, servindo como evidência preliminar para a proposta de transformação desenvolvida nesta Tese. Contudo, o REMD difere da RT/RTAD ao empregar um modelo ARIMA parametrizado e mais custoso computacionalmente no componente de alta frequência, em contraste com a RT/RTAD, que busca uma transformação mais universal da série para equalizar a variância e simplificar a detecção subsequente por métodos mais agnósticos e flexíveis.

O estudo de Audibert et al. [2021] propõe transformar séries temporais univariadas em representações multivariadas, utilizando o *Matrix Profile* para extrair informações não locais e alimentando detectores baseados em Redes Neurais Profundas (DNNs). Embora reconheça a importância de características contextuais para anomalias, a abordagem de Audibert et al. [2021] não aborda diretamente a heteroscedasticidade por meio de estabilização da variância, dependendo da robustez das DNNs para lidar com a complexidade dos dados. Em contraste, a RT/RTAD foca na transformação explícita da série para um espaço de variância equalizada, o que permite a aplicação de detectores mais simples e transparentes e evita a complexidade computacional e de interpretabilidade inerente às DNNs.

Dani et al. [2015] propõem o *Adaptive Threshold for Anomaly Detection* (ATAD), um método de detecção de anomalias em séries temporais por meio de limiares adaptativos, segmentando a série e utilizando médias e desvios padrão locais para definir dinamicamente os limites de decisão. Embora reconheça a necessidade de adaptabilidade frente à heteroscedasticidade ao usar estatísticas locais, o ATAD não transforma a série para equalizar a variância globalmente, e sua eficácia depende da sensibilidade da segmentação e da otimização manual de parâmetros como o número de segmentos e o coeficiente de desvio padrão. A RT/RTAD, por sua vez, busca uma transformação mais universal e menos parametrizada, que estabiliza a variabilidade antes da detecção e permite critérios de anomalia mais simples e consistentes em toda a série transformada.

O *Wavelets, Neural Networks and Hilbert Transform* (WANEH), proposto por Kanarachos et al. [2017], é um algoritmo de aprendizado profundo que combina *wavelets* para decomposição multirresolução, redes neurais para aprender o comportamento normal da série e a Transformada de Hilbert para extração de características em sinais residuais. Enquanto o WANEH é uma abordagem complexa baseada em aprendizado que busca reconstruir o comportamento normal e identificar desvios, o que pode implicitamente auxiliar com a variabilidade em diferentes escalas, sua dependência de treinamento extenso e sua menor interpretabilidade o contrastam com a RT/RTAD, que visa transformar a série de forma analítica para um espaço em que anomalias se tornam mais evidentes e podem ser detectadas com abordagens mais simples e menos intensivas em aprendizado.

Sun et al. [2021] propõem o DOW-FLH, um método de pré-processamento *data-driven* para multissazonalidade e heteroscedasticidade, que normaliza resíduos ao dividi-los pela variância de nível de fluxo antes da detecção de anomalias. Embora o DOW-FLH seja uma abordagem de pré-processamento eficaz para lidar com heteroscedasticidade, ele é dependente de agrupamento por dia da semana e níveis de fluxo, o que pode ser menos flexível do que a decomposição agnóstica, como a CEEMD da RT/RTAD, para capturar nuances da variância variável em séries mais complexas. A RT/RTAD, por sua vez, oferece uma solução integrada que transforma diretamente a série para equalizar a variância, permitindo uma detecção mais direta e adaptável.

A Tabela III.1 resume os principais atributos do trabalho proposto e compara os trabalhos supracitados: transformação como solução independente, uso de decomposição, foco em heteroscedasticidade, foco em anomalias contextuais e detector de anomalias. *Transformação como solução* indica se a transformação é proposta como contribuição independente e reutilizável, e *Uso de decomposição* refere-se ao emprego de técnicas de decomposição. *Foco em heteroscedasticidade* marca os trabalhos que tratam explicitamente da variância não constante, enquanto *Foco em anomalias contextuais* identifica aqueles voltados à detecção de eventos dependentes do comportamento local. Por fim, *Detector de anomalias* diferencia métodos completos de detecção daqueles que apenas transformam a série, permitindo o uso posterior de detectores externos.

Os artigos reunidos na Tabela III.1 ilustram as diversas abordagens na literatura que endereçam a detecção de anomalias em séries temporais, refletindo os caminhos principais de pesquisa aqui identificados. O REMD [Souza et al., 2024], um precursor conceitual para a RT e com autoria compartilhada, exemplifica uma abordagem híbrida que combina decomposição de séries temporais, associada às transformações para estabilização de dispersão, com modelagem estatística para refinar a detecção. A contribuição de Audibert et al. [2021] situa-se na interseção das transformações de dados, utilizando o *Matrix Profile* para gerar informações contextuais, e o emprego de detectores modernos baseados em arquiteturas complexas de aprendizado profundo. O ATAD [Dani et al., 2015] representa as estratégias de adaptação dinâmica de limiares baseadas em estatísticas

Tabela III.1: Resumo comparativo dos trabalhos mais semelhantes à RT/RTAD.

Atributo	**REMD** [Souza et al., 2024]	**Audibert** [Audibert et al., 2021]	**ATAD** [Dani et al., 2015]	**WANEH** [Kanarachos et al., 2017]	**Sun** [Sun et al., 2021]	**RT/RTAD**
Transformação como solução		X			X	X
Uso de decomposição	X			X	X	X
Foco em heteroscedasticidade			X		X	X
Foco em anomalias contextuais		X	X	X		X
Detector de anomalias	X		X	X		X

locais, buscando contornar a heteroscedasticidade por meio de modelagem estatística adaptativa, sem necessariamente uma transformação global. O WANEH [Kanarachos et al., 2017] é um exemplo robusto de como a decomposição multiescala, via *wavelets*, pode ser integrada a modelos de aprendizado profundo para uma detecção abrangente. Finalmente, Sun et al. [2021] destacam-se por propor métodos de pré-processamento explicitamente voltados à estabilização da variância em séries multissazonais, alinhando-se diretamente com o foco em transformações de dispersão.

Essa comparação evidencia que, embora alguns trabalhos compartilhem aspectos da abordagem aqui proposta, apenas a RT/RTAD combina a transformação como solução independente, o tratamento explícito da heteroscedasticidade e a ênfase em anomalias contextuais. Ao converter anomalias contextuais em globais por meio da equalização da variabilidade local, a RT/RTAD oferece um diferencial conceitual frente às abordagens anteriores, representando uma contribuição original para a detecção de anomalias em séries temporais heteroscedásticas.

Capítulo IV Metodologia

A RT é uma transformação projetada para modificar séries temporais heteroscedásticas, visando realçar anomalias, especialmente as de natureza contextual. Essa estratégia é particularmente adequada para séries com variância não constante e fundamenta-se na decomposição da série original em seus componentes de frequência. Para tal, a RT inicia com a aplicação da decomposição do modo empírico complementar (CEEMD) para gerar as funções de modo intrínseco (IMF).

A transformação é realizada com as componentes de alta frequência, que são selecionadas a partir de um limiar definido pela rugosidade das IMF. Um ponto-chave da transformação é o uso da dispersão local para considerar a volatilidade da série, contribuindo para destacar as anomalias contextuais. As detecções no *Resilient Transformation Anomaly Detector* (RTAD) consistem na utilização da série transformada pela RT, aliada à decisão por limiar gaussiano.

Neste capítulo, são apresentadas todas as etapas de execução da RT, bem como a utilização da série transformada para detecção de anomalias (RTAD).

IV.1 Transformação Resiliente

A RT consiste em seis etapas: (i) decomposição por EMD; (ii) seleção de componentes; (iii) composição das IMF; (iv) diferenciação; (v) cálculo da dispersão local; e (vi) transformação da série final. As principais notações empregadas ao longo deste capítulo estão detalhadas na Tabela IV.1.

Para facilitar a compreensão do processo da RT, a Figura IV.1 apresenta um fluxograma que sumariza as etapas discutidas neste capítulo. O diagrama ilustra as etapas sequenciais da RT, destacando a decomposição inicial da série original Y_t e a subsequente ramificação em duas trilhas de processamento paralelas: uma para a composição e diferenciação das IMF de alta frequência (culminando em δ_t) e outra para o cálculo da dispersão local da série sem tendência (resultando em ψ_t). Ambas as trilhas convergem na etapa final para gerar a série transformada α_t, pronta para a detecção de anomalias. A figura estabelece uma correspondência visual direta com as etapas descritas textualmente e com os procedimentos detalhados no Algoritmo 1.

O Algoritmo 1 resume os procedimentos envolvidos na RT, detalhando a sequência lógica e as operações matemáticas desde a decomposição inicial até a geração da série transformada. As linhas do algoritmo estão diretamente correlacionadas com as etapas descritas textualmente e ilustradas

Tabela IV.1: Símbolos utilizados na RT e no RTAD.

Símbolo	Descrição
n	Número de elementos na série temporal
Y_t	Série temporal de entrada, onde $1 \leq t \leq n$
$\iota_t^{1:k}$	Conjunto de IMF do CEEMD, da 1ª à k-ésima
k	Número de IMF geradas pelo CEEMD
ρ_t	Componente residual do CEEMD
γ_j	Rugosidade da IMF j, onde $1 \leq j \leq k$
d	Número de IMF selecionadas (limite inferior)
σ_t	Soma das IMF de alta frequência
δ_t	Diferença de primeira ordem de σ_t
ξ_t	Série sem tendência
w	Tamanho da janela
ψ_t	Série de dispersão local calculada em janela w
α_t	Série final transformada (saída da RT)
L_S	Limiar superior para detecção de anomalias
L_I	Limiar inferior para detecção de anomalias
λ	Multiplicador do desvio padrão para definição dos limiares
$noise.amp$	Amplitude do ruído branco adicionado no CEEMD
$trials$	Número de execuções de ensaios no CEEMD

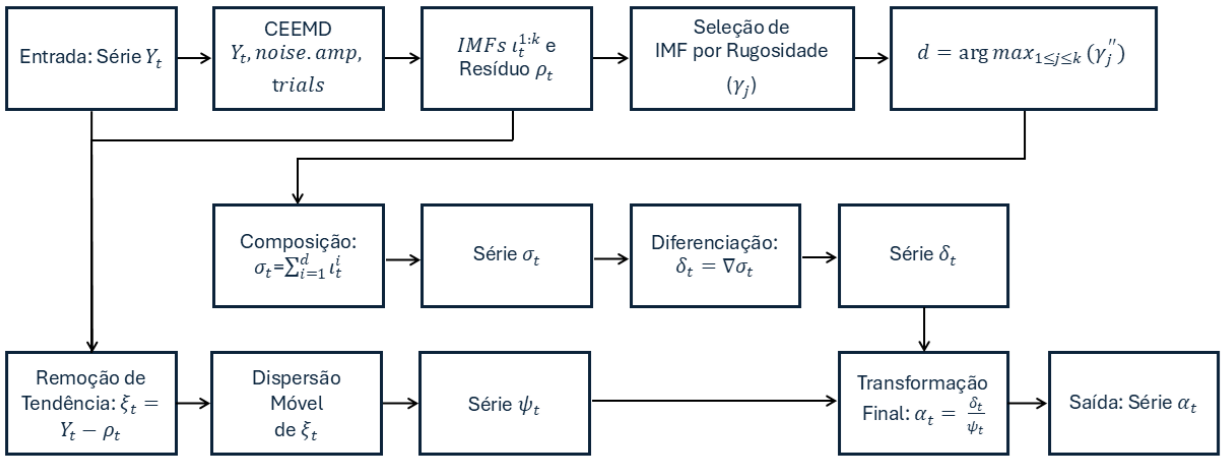


Figura IV.1: Fluxograma esquemático das etapas da Transformação Resiliente (RT), desde a decomposição da série temporal original até a obtenção da série transformada final.

no fluxograma, utilizando as notações definidas na Tabela IV.1.

Algoritmo 1 Transformação Resiliente (RT)

Entrada: Série temporal Y_t, parâmetros $noise.amp$, $trials$, w
Saída: Série transformada α_t
Função RT(Y_t, $noise.amp$, $trials$, w)
    $\iota_t^{1:k}, \rho_t \leftarrow \text{CEEMD}(Y_t, noise.amp, trials)$ ▷ Decomposição
    $\gamma_j \leftarrow \text{rugosidade}(\iota_t^j)$ ▷ Seleção de componentes
    $d \leftarrow \arg\max_{1\leq j\leq k}(\gamma_j'')$
    $\sigma_t \leftarrow \sum_{i=1}^{d} \iota_t^i$ ▷ Composição das IMF
    $\delta_t \leftarrow \nabla\sigma_t$ ▷ Diferenciação
    $\xi_t \leftarrow Y_t - \rho_t$ ▷ Remoção da tendência
    $\psi_t \leftarrow \text{DispersaoLocal}(\xi_{t-w+1:t})$ ▷ Cálculo da dispersão local
    $\alpha_t \leftarrow \dfrac{\delta_t}{\psi_t}$ ▷ Transformação final
    retorne α_t

Decomposição

A etapa de decomposição (linha 4 do Algoritmo 1) é a base da RT, em que a série temporal de entrada Y_t é processada pelo CEEMD. Essa técnica decompõe a série original em um conjunto de componentes oscilatórias intrínsecas, as IMF, e um componente residual (vide Seção II.2). O princípio do CEEMD é representar a série como uma combinação de sinais com diferentes escalas de variação, permitindo a separação entre flutuações de alta frequência (associadas a oscilações locais abruptas) e componentes mais suaves e estruturais. Tal separação é crucial no contexto da RT, pois possibilita isolar as IMF mais sensíveis a anomalias.

Considere uma série temporal Y_t na forma $\langle y_1, y_2, y_3, \ldots, y_n \rangle$, em que n é o número de observações. O primeiro passo é aplicar o CEEMD à série Y_t. Esse processo exige dois parâmetros de entrada adicionais: *noise.amp* e *trials*. O parâmetro *noise.amp* define a amplitude relativa do ruído branco adicionado à série original durante a decomposição. O parâmetro *trials* especifica o número de execuções de decomposição independentes, cada uma utilizando diferentes amostras de ruído.

Neste trabalho, são utilizados valores padrão de $noise.amp = 0{,}001$ e $trials = 5$ para a aplicação do CEEMD. Uma baixa amplitude de ruído é escolhida para minimizar a interferência com os componentes reais da série, garantindo que o processo de decomposição introduza apenas o ruído necessário para evitar a sobreposição de escalas. O número reduzido de execuções representa um equilíbrio entre a estabilidade da decomposição e o custo computacional, proporcionando uma suavização suficiente dos efeitos de ruído sem um tempo de processamento excessivo. Ambos os parâmetros podem ser ajustados durante a execução da RT.

A saída resultante é uma lista de vários elementos, incluindo as séries de IMF geradas. O conjunto de IMF gerado pelo CEEMD é denotado por $\iota_t^{1:k}$, compreendendo k séries temporais. Como característica do processo de peneiramento do CEEMD, a primeira IMF (ι_t^1) possui a maior frequência, e a frequência diminui progressivamente até o último componente (ι_t^k). O valor de k varia conforme a série analisada e é determinado principalmente pela complexidade do sinal.

Seleção de Componentes

A etapa de seleção de componentes (linha 5 do Algoritmo 1) visa identificar quais IMF devem ser incluídas na série reconstruída para realçar as flutuações mais relevantes. A determinação do ponto de divisão das IMF (denotado por d) começa com o cálculo da rugosidade de cada uma delas. Para cada IMF ι_t^j, um valor escalar γ_j é calculado, representando sua rugosidade, em que $1 \leq j \leq k$. A rugosidade γ_j é calculada de acordo com a Equação II.3, apresentada na Seção II.1.5.

Seja $\boldsymbol{\gamma} = \langle \gamma_1, \gamma_2, \ldots, \gamma_k \rangle$ a sequência de valores de rugosidade das k IMF. Em geral, essa sequência apresenta uma tendência decrescente à medida que se avança dos componentes de alta para os

de baixa frequência. O índice d é definido como o ponto de máxima curvatura dessa sequência, isto é, o valor de j que maximiza γ''_j, conforme a equação apresentada na Seção II.1.6.

O índice d representa o limiar a partir do qual as IMF predominantemente contêm oscilações mais suaves e estruturais. Portanto, apenas os primeiros d componentes são utilizados na reconstrução da série transformada. A Figura IV.2 ilustra o comportamento típico dos valores de rugosidade das IMF acumuladas e o destaque do ponto de curvatura máxima (em vermelho), pertinente para o cálculo de d no cenário geral. Contudo, em situações específicas nas quais o CEEMD não gera mais do que três IMF ($k \leq 3$), o cálculo da curvatura máxima torna-se inviável. Nestes casos excepcionais, o valor de d é, por definição, igual a k, o que implica que todas as IMF produzidas são incluídas na etapa subsequente de composição.

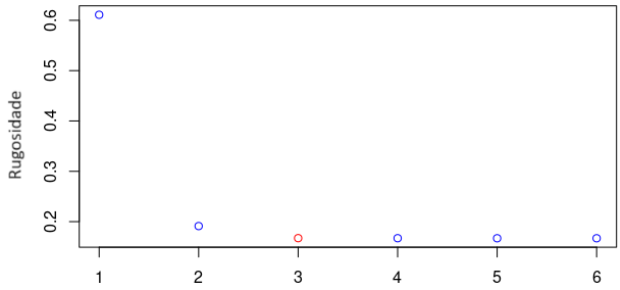


Figura IV.2: Valores de rugosidade das IMF. Destacado em vermelho o ponto de curvatura máxima (d).

Composição das IMF de alta frequência

O próximo passo na RT é a composição das IMF selecionadas (linha 7 do Algoritmo 1), com o objetivo de reconstruir uma versão filtrada da série original que enfatize oscilações locais de alta frequência associadas à ocorrência de anomalias. A série σ_t é obtida somando as primeiras d IMF, que são os componentes de maior frequência, conforme identificado na etapa anterior e formalmente na Equação IV.1:

$$\sigma_t = \sum_{i=1}^{d} \iota_t^i \tag{IV.1}$$

É importante ressaltar que, na teoria do CEEMD e suas variantes, a soma de todas as IMF geradas e o componente residual ρ_t reproduzem a série original Y_t, ou seja, $Y_t = \sum_{i=1}^{k} \iota_t^i + \rho_t$. Consequentemente, se todas as IMF são selecionadas para a composição (isto é, quando $d = k$), a série σ_t será matematicamente $\sigma_t = Y_t - \rho_t$. Este cenário de $d = k$ ocorre onde o número de IMF produzidas pelo CEEMD é limitado a três ou menos, como explicado na etapa anterior.

Diferenciação

Em seguida, a série composta passa por diferenciação (linha 8 do Algoritmo 1) para realçar transições de curto prazo e reduzir a persistência temporal. Para isso, é computada a diferença de primeira ordem de σ_t. Diferentemente das etapas anteriores, em que as operações envolviam séries distintas, a diferenciação opera sobre termos sucessivos de σ_t. A série resultante é δ_t, conforme a Equação II.2 (ver Seção II.1.2). Como a diferenciação reduz o comprimento da série em uma unidade, um novo termo é inserido na posição $t = 1$ (δ_1), com valor igual ao segundo termo (δ_2), para preservar o comprimento original da série.

Cálculo da Dispersão Local

A etapa subsequente envolve o cálculo da dispersão local (linha 9 do Algoritmo 1), essencial para padronizar a série diferenciada. Inicialmente, obtém-se a série sem tendência ξ_t, subtraindo-se o componente residual ρ_t da série original Y_t, conforme a Equação IV.2. Na decomposição empírica, o componente residual ρ_t concentra as oscilações de baixa frequência e é interpretado como a tendência da série original.

$$\xi_t = Y_t - \rho_t \tag{IV.2}$$

Uma série de dispersão local ψ_t é então calculada utilizando uma janela deslizante de tamanho fixo. Adota-se uma janela alinhada à direita, o que significa que o valor de ψ_t na posição t reflete a dispersão dos valores precedentes, na forma $\langle \xi_{t-w+1}, \xi_{t-w+2}, \ldots, \xi_t \rangle$, em que w denota o tamanho da janela.

No início da série (para $t < w$), um número reduzido de observações é utilizado, evitando a perda de dados e permitindo o cálculo mesmo quando a janela completa ainda não está disponível. Este trabalho utiliza o desvio padrão como medida de dispersão, embora outras alternativas também possam ser empregadas, como discutido na Seção V.2.4.

Um tamanho de janela de 30 observações ($w = 30$) é adotado neste trabalho, pois esse valor é amplamente utilizado em estudos de séries temporais para equilibrar a estabilidade estatística e a sensibilidade local [Shumway and Stoffer, 2017; Hamilton, 1994]. Tradicionalmente, esse tamanho é considerado suficiente para que as estimativas da amostra se aproximem das propriedades assintóticas sob o Teorema do Limite Central [Moore et al., 2014], o que o tornou um padrão em muitas abordagens baseadas em janelas locais. Contudo, é fundamental reconhecer que a velocidade de convergência à normalidade, conforme previsto pelo Teorema do Limite Central, pode ser significativamente afetada por características da distribuição original dos dados, tais como caudas pesadas, assimetria ou a própria heteroscedasticidade. Portanto, o desempenho da transformação pode va-

riar substancialmente com w, sendo aconselhável que esse parâmetro seja ajustado considerando as características específicas da série e os objetivos da aplicação.

Transformação da Série Final

Finalmente, a série transformada α_t é obtida padronizando a série diferenciada δ_t em relação à dispersão local ψ_t (linha 11 do Algoritmo 1). A Equação IV.3 define α_t como a razão de δ_t para ψ_t. Essa operação produz uma série com amplitude aproximadamente constante ao longo do tempo, na qual eventos são realçados por seus desvios em relação à variabilidade local esperada.

É importante notar que, dependendo da função de dispersão e do tamanho da janela local, especialmente para as primeiras observações da série, quando a janela não está completamente preenchida, o cálculo de ψ_t pode resultar em valores indefinidos (NA), por exemplo, se a medida de dispersão não puder ser computada para um subconjunto com poucas observações. Nesses casos, os elementos correspondentes na série transformada α_t também apresentarão valores NA.

$$\alpha_t = \frac{\delta_t}{\psi_t} \tag{IV.3}$$

A Figura IV.3 apresenta um exemplo das séries δ_t, ψ_t e α_t. É possível notar que δ_t já evidencia as anomalias; porém, quando é dividida por ψ_t, que mostra três níveis distintos de variância ao longo da série, as anomalias evidenciadas passam a ter amplitudes semelhantes. Em δ_t, uma delas acaba tendo menor amplitude por estar em um trecho de menor variância. A série α_t é bastante homogênea, apresentando variação expressiva apenas nas anomalias.

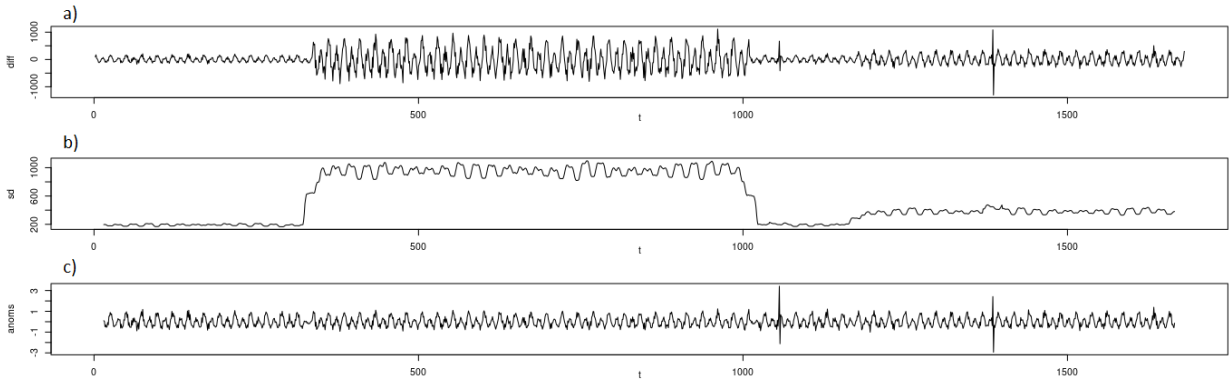


Figura IV.3: Séries a) δ_t, b) ψ_t e c) α_t.

IV.2 Detecção de Anomalias com RTAD

A série transformada α_t, resultante da aplicação da RT conforme detalhado na seção anterior, é o ponto de partida para a detecção de anomalias. O método completo, aqui referido como RTAD, compreende a utilização de α_t, seguida pela regularização L1, aplicação de limiares estatísticos e

um filtro de pós-processamento para a identificação final das anomalias.

Para refinar a detecção de anomalias e enfatizar os eventos significativos, a série α_t é inicialmente submetida a um processo de regularização L1, que consiste em considerar o valor absoluto de seus elementos. Esta operação transforma todos os valores em não-negativos, permitindo que anomalias que originalmente se manifestavam como desvios negativos sejam tratadas de forma equivalente às positivas, focando na magnitude do desvio. O objetivo é realçar os picos correspondentes a anomalias e atenuar o ruído de fundo, facilitando a distinção entre variações normais e eventos anômalos.

Sobre a série regularizada, o processo de detecção de anomalias é baseado em limiares gaussianos. O limiar superior (L_S) é definido como a média dos valores da série regularizada mais λ desvios padrão. Adota-se $\lambda = 3$. Pontos na série regularizada que excedem L_S são classificados como anomalias potenciais. Neste contexto, o limiar inferior (L_I), embora calculável, torna-se irrelevante para a detecção de anomalias, pois todos os valores da série regularizada são não-negativos, e valores muito baixos (próximos de zero) indicam ausência de anomalia.

Adicionalmente, para evitar múltiplas detecções do mesmo evento em sequência, é aplicado um filtro de pós-processamento que retém apenas a primeira ocorrência em cada grupo de detecções consecutivas. Essa etapa garante que o conjunto de eventos detectados reflita a ocorrência isolada de eventos ao longo do tempo.

IV.3 Implementação

O *framework* Harbinger é utilizado neste trabalho para executar todos os métodos e métricas necessários para a implementação, avaliação e análise gráfica dos resultados. O Harbinger é um *framework* para integração e análise de métodos de detecção de eventos. Ele fornece vários métodos de última geração para traçar e avaliar detecções de eventos [Ogasawara et al., 2025a; Salles et al., 2020]. O Harbinger está disponível como um pacote R[1] e é constantemente atualizado, inclusive com a introdução de novos métodos. O RTAD foi implementado no código `hanr_rtad`, em R, e está disponível no GitHub[2].

Para a execução do código, o único parâmetro de entrada necessário é a série temporal cujas anomalias se deseja detectar. Alguns parâmetros são opcionais, visto que o padrão adotado atende com eficiência à proposta de detectar anomalias de maneira abrangente, para séries temporais de naturezas variadas. Porém, esses parâmetros podem ser alterados conforme a necessidade do usuário. É possível escolher o tamanho da janela e os parâmetros do CEEMD. Os parâmetros da RT que são configuráveis no método RTAD são resumidos na Tabela IV.2.

[1] Harbinger R *package*: `https://cefet-rj-dal.github.io/harbinger/`

[2] `https://github.com/cefet-rj-dal/harbinger/blob/master/R/hanr_rtad.R`

Tabela IV.2: Parâmetros opcionais da RT no contexto do RTAD.

Parâmetro	**Explicação**	**Default**
noise.amp	Parâmetro do CEEMD que determina a amplitude de ruído branco	0,001
trials	Parâmetro do CEEMD que determina o número de execuções	5
w	Tamanho da janela para cálculo do desvio padrão local	30

IV.4 Exemplo ilustrativo

Para elucidar a aplicação prática da RT e do RTAD, esta seção apresenta um exemplo ilustrativo detalhado. A série temporal utilizada é um trecho de uma das séries heterocedásticas empregadas na avaliação experimental, conforme detalhado na Seção V.1.1. Este segmento é caracterizado por um trecho inicial de baixa variância com uma leve tendência decrescente, seguido por um trecho de variância acentuadamente mais alta. Existem nessa série cinco anomalias pontuais: quatro no período de baixa variância e uma no de alta variância, permitindo demonstrar a capacidade do RTAD de lidar com condições heterocedásticas. A Figura IV.4 exibe a série temporal de entrada, Y_t.

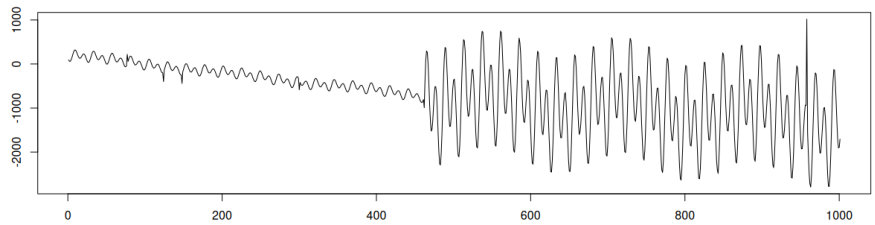


Figura IV.4: Série temporal Y_t de entrada para o exemplo ilustrativo, com períodos de baixa e alta variância e cinco anomalias.

O processo da RT inicia com a decomposição da série Y_t pelo CEEMD (linha 4 do Algoritmo 1), gerando o conjunto de IMF $\iota_t^{1:k}$ e um componente residual ρ_t. A Figura IV.5 apresenta as IMF e A Figura IV.6 o residual gerados para a série de exemplo.

A seleção dos componentes de alta frequência é realizada pela análise da rugosidade de cada IMF (linha 5), identificando o ponto de máxima curvatura, d, que neste exemplo foi 2. A Figura IV.7 ilustra a variação da rugosidade e o ponto de corte para a seleção das IMF. As primeiras d IMF são então somadas para formar a série σ_t (linha 7), conforme exibido na Figura IV.8.

Subsequentemente, a série σ_t é diferenciada para obter δ_t (linha 8). A Figura IV.9 ilustra a série δ_t.

Em paralelo, a série sem tendência ξ_t é calculada subtraindo o residual ρ_t de Y_t. A Figura IV.10 A Figura IV.10 apresenta esta série.

Com base em ξ_t, computa-se a série de dispersão local ψ_t utilizando uma janela deslizante de $w = 30$ (linha 9). A Figura IV.11 apresenta a série ψ_t.

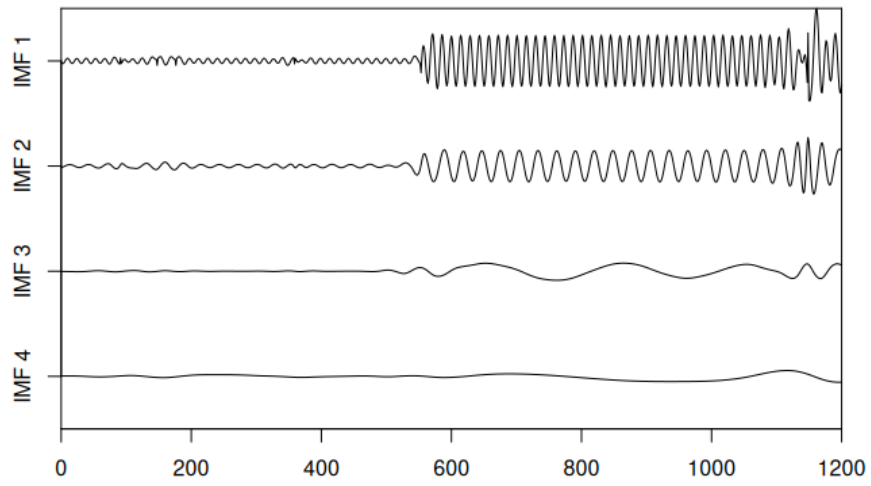


Figura IV.5: IMF resultantes da decomposição CEEMD da série temporal Y_t ($\iota_t^{1:k}$).

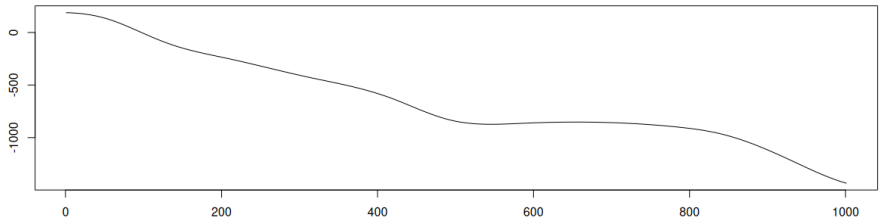


Figura IV.6: Componente residual ρ_t resultante da decomposição CEEMD da série temporal Y_t.

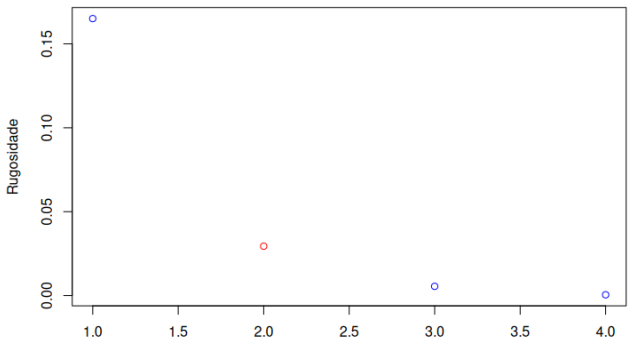


Figura IV.7: Rugosidade das IMF da série de exemplo, com o ponto de máxima curvatura ($d = X$) destacado em vermelho.

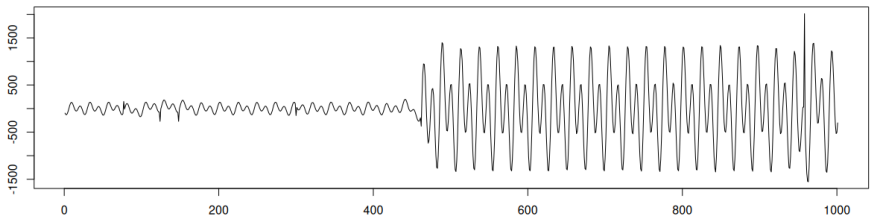


Figura IV.8: Série σ_t, obtida pela soma das IMF de alta frequência selecionadas, capturando as flutuações rápidas da série de entrada.

Finalmente, a série transformada α_t (linha 11) é gerada pela razão δ_t/ψ_t (Equação IV.3). A Figura IV.12 ilustra a série α_t, demonstrando como a padronização pela dispersão local homogeneíza a amplitude dos picos anômalos em um contexto de variância não estacionária.

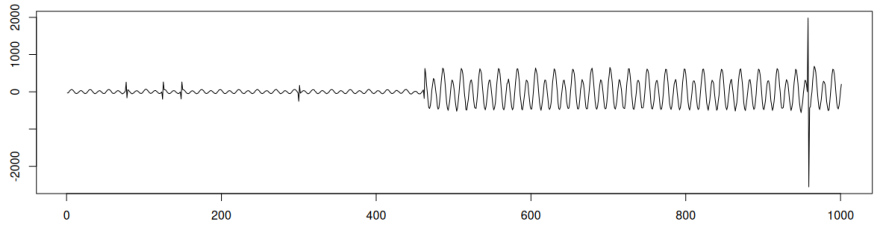


Figura IV.9: Série δ_t, resultante da diferenciação de σ_t.

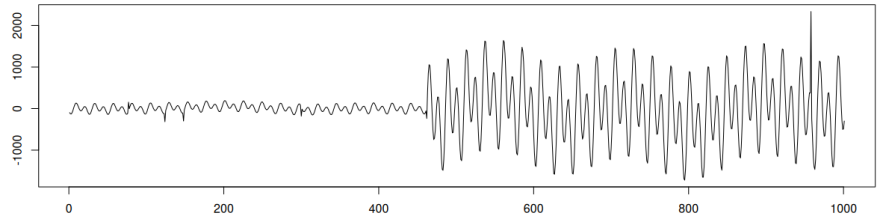


Figura IV.10: Série ξ_t, obtida a partir de Y_t após a remoção do componente residual ρ_t, representando a série sem tendência.

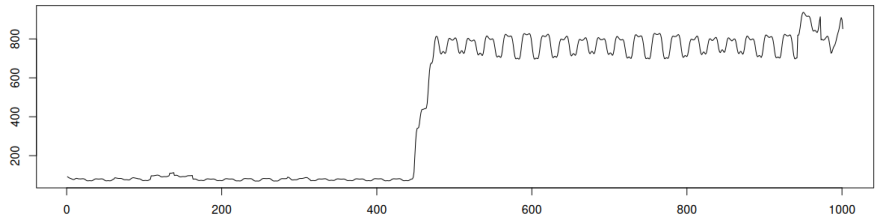


Figura IV.11: Série ψ_t, representando a dispersão local calculada na série sem tendência ξ_t.

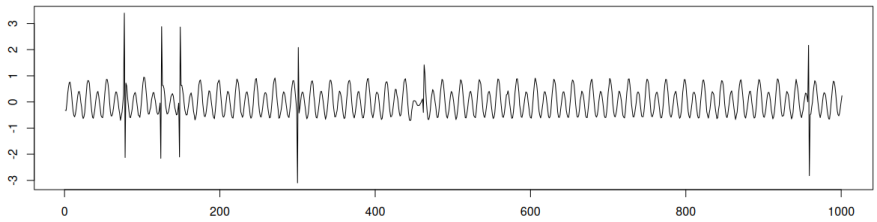


Figura IV.12: Série α_t, a série final transformada pela RT, onde as anomalias são realçadas com amplitude homogênea, facilitando a detecção.

Para a detecção de anomalias, aplica-se um critério de limiares gaussianos à série α_t. Um filtro de pós-processamento é aplicado para identificar apenas a primeira ocorrência em grupos consecutivos de detecções. A Figura IV.13 apresenta a série α_t regularizada juntamente com os limiares e as anomalias detectadas. O exemplo demonstra que o RTAD consegue identificar com sucesso as anomalias tanto nos segmentos de baixa quanto de alta variância.

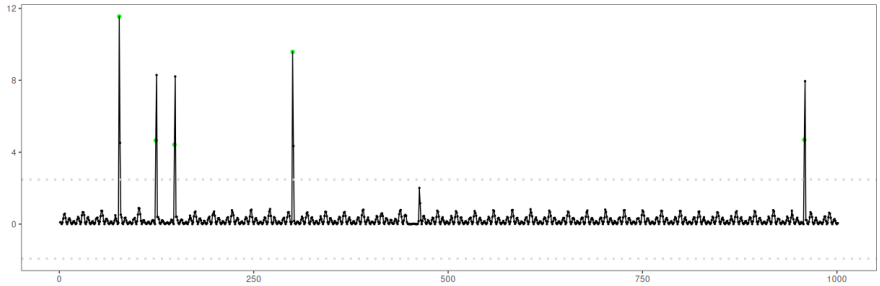


Figura IV.13: Série transformada α_t com os limiares de detecção (L_S e L_I) e os pontos de anomalia detectados pelo RTAD, ilustrando a identificação precisa das anomalias.

Capítulo V Avaliação Experimental

Este capítulo apresenta a avaliação experimental conduzida para examinar a influência da utilização da RT na detecção de anomalias em séries temporais. A Seção V.1 descreve os elementos que dão suporte aos experimentos. A Seção V.2 detalha os resultados da avaliação do impacto da RT como etapa de pré-processamento capaz de aprimorar o desempenho de métodos de detecção existentes e do desempenho do RTAD frente a outros métodos de detecção de eventos, além de explorar a sensibilidade do método proposto (RTAD) a diferentes configurações. A Seção V.3 sumariza e discute os resultados.

V.1 Configuração Experimental

Os experimentos são conduzidos utilizando o *framework* Harbinger [Ogasawara et al., 2025a; Salles et al., 2020], que fornece implementações padronizadas de métodos de detecção e métricas de avaliação. Os testes foram executados em uma máquina com as seguintes configurações de *hardware*: processador Intel Xeon w3-2423, 512 GB de RAM, 12 núcleos e sistema operacional Ubuntu 22.04 LTS.

São considerados nove modelos de linha de base, avaliados por três métricas primárias (F1, precisão e revocação), com as comparações estatísticas realizadas utilizando o teste de Wilcoxon pareado e a estimativa do tamanho do efeito.

V.1.1 Conjunto de Dados

Os experimentos deste estudo utilizam séries temporais extraídas do conjunto de dados *Yahoo! S5*, amplamente reconhecido e empregado na avaliação de métodos de detecção de anomalias [Thill et al., 2017; Yin et al., 2022; Buda et al., 2018; Garg et al., 2020]. Este conjunto, originalmente fornecido pela Verizon Media Webscope para uso não comercial, é acessado no presente trabalho por meio do pacote UnitED[1] [Lima et al., 2025].

O UnitED (*A Unified Time Series Event Detection Repository*) constitui um repositório curado de conjuntos de dados rotulados, concebido para o desenvolvimento e a avaliação consistente de métodos de detecção de eventos em séries temporais. O *Yahoo! S5* compreende mais de 300 séries

[1] `https://github.com/cefet-rj-dal/united`

temporais, incluindo tanto dados reais quanto sintéticos, todas com eventos rotulados (anomalias e pontos de mudança, quando presentes). Neste estudo, foram selecionados dois subconjuntos específicos do *Yahoo! S5*: o Benchmark A3 e o Benchmark A4.

Para fins da avaliação experimental, as séries foram divididas em subconjuntos homoscedásticos e heteroscedásticos. Esta classificação baseou-se no teste formal de heteroscedasticidade de Breusch–Pagan (vide Seção II.1.4). Todas as 100 séries do Benchmark A3 foram classificadas como homoscedásticas, formando o conjunto correspondente. Do Benchmark A4, 74 séries foram identificadas como heteroscedásticas e selecionadas para formar o grupo heteroscedástico.

Em termos de características das anomalias, o conjunto homoscedástico (Benchmark A3) totaliza 943 anomalias, com uma média de $9,43$ anomalias por série. O conjunto heteroscedástico (Benchmark A4) possui 635 anomalias, apresentando uma média de $8,58$ anomalias por série. Em ambos os conjuntos, a quantidade de anomalias por série varia de 1 a 16. A Figura V.1 mostra exemplos de séries dos dois grupos.

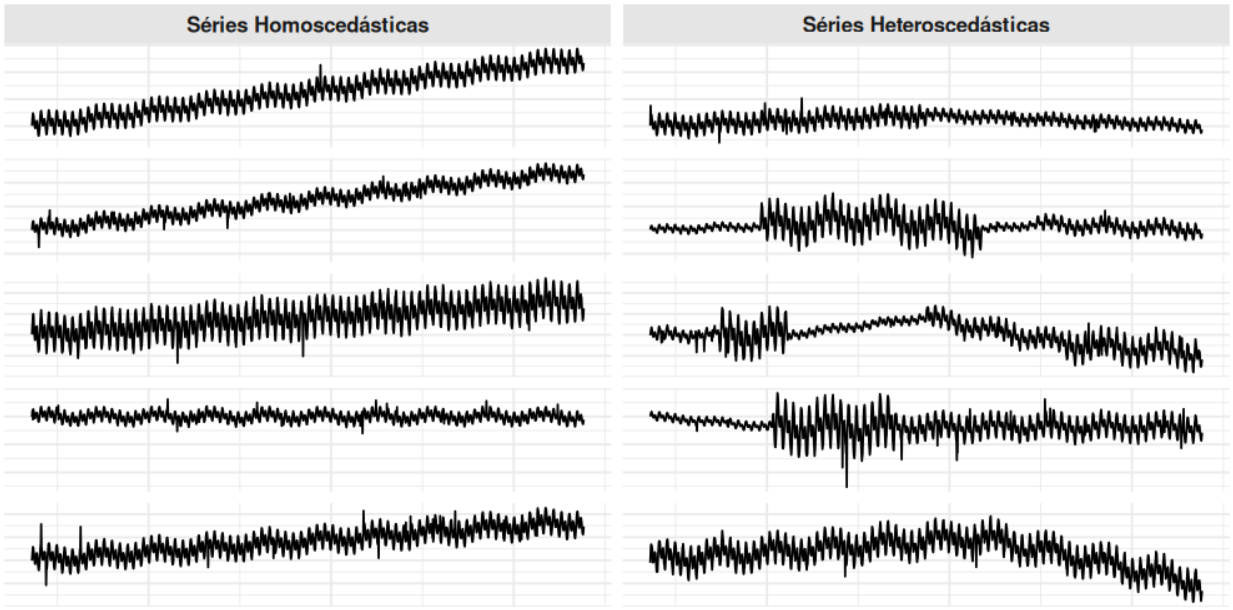


Figura V.1: Séries temporais de índices 10, 20, 30, 40 e 50, extraídas de dois conjuntos de dados: séries homoscedásticas (coluna da esquerda) e séries heteroscedásticas (coluna da direita).

V.1.2 Medidas de Avaliação

Três métricas são utilizadas para avaliar o desempenho dos métodos de detecção: F1, precisão e revocação. Conforme detalhado na Seção II.4, altos valores de F1 exigem resultados fortes em ambas as métricas e são frequentemente usados como indicador resumido da qualidade geral da detecção [Saito and Rehmsmeier, 2015].

É importante notar que, no cálculo da métrica F1, quando tanto a precisão quanto a revocação resultam em zero, o F1 torna-se indeterminado. Nesses casos, os valores correspondentes foram

desconsiderados da média final, assegurando que apenas resultados válidos contribuam para o cômputo da média. Ou seja, se um método não detectar anomalias corretamente em algumas séries do conjunto, a média de F1 é calculada apenas sobre as séries com valores válidos para precisão *ou* revocação. Cabe complementar que esse cenário ocorre em cerca de 7% das séries originais e em menos de 1% das séries transformadas por RT, considerando todos os métodos de detecção utilizados.

V.1.3 Modelos de Linha de Base

Nove métodos de detecção de anomalias de última geração são considerados e utilizados como linhas de base para avaliar o desempenho da RT. Todos os métodos estão disponíveis no *framework* Harbinger com parâmetros padrão fornecidos pela biblioteca. As técnicas de modelagem específicas são descritas a seguir e organizadas em grupos:

Modelos de transformação espectral: neste grupo, são exploradas técnicas que decompõem a série temporal em diferentes domínios para identificar padrões subjacentes e modelar seu comportamento. A *Transformada Rápida de Fourier (FFT)* realiza a modelagem da série no domínio da frequência, decompondo-a em componentes harmônicas. Após a análise espectral, as frequências de menor relevância são eliminadas, e a série é reconstruída por transformada inversa [Cooley and Tukey, 1965], permitindo isolar o comportamento esperado. A abordagem baseada em *Transformada Wavelet* decompõe a série temporal em múltiplas escalas de variação por meio da Transformada Discreta com Sobreposição. Os coeficientes de detalhe resultantes de diferentes níveis de decomposição são agregados e utilizados para representar o comportamento modelado da série [Percival and Walden, 2000], capturando características em diferentes resoluções temporais.

Modelos estatísticos tradicionais: esta categoria foca abordagens estatísticas amplamente estabelecidas para a modelagem de séries temporais. O modelo *AutoRegressive Integrated Moving Average (ARIMA)* modela a série com base em relações lineares entre observações passadas e seus próprios erros. A estrutura do modelo é automaticamente definida a partir de critérios estatísticos e testes de estacionariedade [Hyndman and Khandakar, 2008], sendo uma ferramenta fundamental para séries univariadas. Complementarmente, o modelo *Generalized Autoregressive Conditional Heteroskedasticity (GARCH)* é empregado quando a variância dos erros é condicional e evolui ao longo do tempo. A série modelada, nesse caso, corresponde aos resíduos padronizados obtidos após o ajuste, refletindo variações inesperadas na volatilidade condicional. Essa abordagem é particularmente adequada para séries heteroscedásticas, permitindo capturar eventos associados a aumentos abruptos de variância [Bollerslev, 1986].

Modelos empíricos e híbridos: este grupo reúne métodos que combinam diferentes princípios para modelar a série temporal, seja por meio de decomposição baseada em dados, seja por

abordagens híbridas. O método *Forward and Backward Inertial Anomaly Detector (FBIAD)* realiza a modelagem com base na inércia local de janelas deslizantes, analisadas tanto no sentido direto quanto no reverso da série. Ele calcula o desvio entre o valor observado e a média de sua vizinhança anterior (forward) e posterior (backward), combinando essas perspectivas para identificar anomalias que se destacam em relação ao contexto local em ambas as direções [Lima et al., 2022]. Por sua vez, o *Refined Empirical Mode Decomposition (REMD)* adota uma abordagem híbrida avançada, combinando a decomposição do modo empírico (EMD) com a modelagem ARIMA. A série é inicialmente decomposta em componentes oscilatórios (IMF), dos quais se extraem as frequências mais altas por meio de análise de rugosidade cumulativa. Um modelo ARIMA é ajustado a esses componentes, e a série modelada é composta pelas previsões resultantes [Souza et al., 2024].

Modelos de aprendizado de máquina: um conjunto de técnicas baseadas em aprendizado de máquina foi utilizado para capturar padrões complexos e não lineares nas séries temporais. O modelo *Support Vector Machine (SVM)* aplica máquinas de vetor de suporte para regressão, ajustando uma função que se aproxima da série com tolerância a desvios dentro de uma margem. A modelagem é realizada a partir de janelas deslizantes que capturam a estrutura temporal local [Smola and Schölkopf, 2004]. A *Multi-Layer Perceptron (MLP)*, uma rede neural, realiza a modelagem transformando a série em janelas deslizantes, nas quais sequências de valores anteriores predizem o próximo. A rede aprende os padrões sequenciais e gera a série modelada com base em suas previsões [Bishop, 2016]. Para dados sequenciais, o *Long Short-Term Memory (LSTM)*, um tipo de rede neural recorrente, utiliza sua memória de longo prazo para prever o próximo valor a partir de janelas deslizantes, capturando dependências temporais complexas. Anomalias são detectadas pelos erros de predição, indicando desvios da dinâmica aprendida pela rede.

Entre os cinco trabalhos relacionados apresentados na Tabela III.1, apenas o REMD fornece uma implementação pública. Os demais descrevem seus métodos, mas não liberam código-fonte aberto, o que impede a reprodução direta. No entanto, vários dos modelos de linha de base podem ser considerados representativos dessas abordagens: ARIMA e GARCH capturam limiarização adaptativa como no ATAD; WAVELET e redes neurais aproximam o design híbrido do WANEH; ARIMA, WAVELET e GARCH refletem aspectos do pré-processamento sazonal e heteroscedástico de Sun et al. [2021]; e os detectores de aprendizado profundo são comparáveis aos empregados por Audibert et al. [2021]. Assim, embora apenas o REMD pudesse ser incluído diretamente, a avaliação experimental ainda abrange as principais dimensões metodológicas dos outros trabalhos relacionados.

Independentemente da técnica utilizada, todos os métodos seguem a mesma lógica de detecção implementada no Harbinger. Cada modelo ajusta a série temporal, calcula a série residual (diferença ponto a ponto entre os valores observados e estimados) e aplica uma métrica baseada na distância

quadrática L_2 (soma dos resíduos ao quadrado). Anomalias preliminares são detectadas com base em um limiar gaussiano, considerando valores que excedem a média em mais ou menos três desvios padrão. Finalmente, um filtro retém apenas a primeira ocorrência de cada grupo de anomalias consecutivas para evitar detecções redundantes do mesmo evento.

V.1.4 Otimização de Hiperparâmetros

A otimização de hiperparâmetros foi realizada para garantir o melhor desempenho e a adequada generalização dos modelos de aprendizado de máquina. O processo de otimização foi conduzido nas séries temporais originais, e os parâmetros resultantes foram aplicados diretamente às respectivas versões transformadas pela RT. Esse procedimento garantiu consistência metodológica e permitiu uma comparação justa do desempenho dos modelos em diferentes representações de dados. As análises nesta seção, portanto, consideram os melhores resultados obtidos após a otimização. A Tabela V.1 apresenta os hiperparâmetros ajustados para os modelos de aprendizado de máquina.

Tabela V.1: Hiperparâmetros variados na otimização dos modelos de aprendizado de máquina.

Modelo	**Parâmetro**	**Valores variáveis**
MLP	**size** (neurônios na camada oculta)	1; 2; . . . ; 10
	decay (regularização por decaimento de peso)	0,0; 0,111; . . . ; 1,0 (passo de 1/9)
SVM	**epsilon** (margem de insensibilidade do SVR)	0,0; 0,1; . . . , 1,0 (passo de 0,1)
	cost (penalidade para erros)	20; 40; 60; 80; 100
	kernel (função *kernel*)	sigmoid (fixo)
LSTM	**epochs** (número de épocas)	2000 (fixo)
Todos	**input_size** (tamanho da janela de entrada)	3; 4; 5; 6; 7

V.1.5 Validação Estatística

Para as análises apresentadas na Seção V.2, emprega-se o teste de Wilcoxon pareado, uma técnica estatística não paramétrica cuja fundamentação e uso para comparar duas amostras pareadas foram detalhados na Seção II.4.1. Adota-se hipótese alternativa bilateral, investigando se as métricas obtidas utilizando RT diferem daquelas sem RT (ou com métodos alternativos, dependendo do contexto da análise). Uma interpretação favorável à proposta desta tese requer que, além de diferença estatisticamente significativa, as métricas obtidas com RT demonstrem superioridade.

Complementarmente à significância estatística, o tamanho do efeito é calculado para quantificar a magnitude da diferença observada entre resultados pareados, conforme a discussão sobre sua importância para a relevância prática na Seção II.4.1. Os testes são realizados de forma pareada para garantir o alinhamento adequado dos resultados.

A interpretação dos achados considera conjuntamente o p-valor e o tamanho do efeito. Diferenças são consideradas estatisticamente significativas para p-valores $\leq 0{,}05$. Neste estudo, o tamanho do

efeito r é calculado como a razão entre o escore padronizado Z (obtido pela padronização Z-Score, vide Seção II.1.1) e a raiz quadrada do tamanho da amostra. O valor de r varia de 0 a quase 1, sendo interpretado conforme os limiares já apresentados na Seção II.4.1: $0{,}1 \leq r < 0{,}3$ (pequeno), $0{,}3 \leq r < 0{,}5$ (médio) e $r \geq 0{,}5$ (grande).

V.2 Experimentos

Os experimentos são conduzidos com os seguintes objetivos: (i) avaliar o impacto da RT em diferentes métodos de detecção de anomalias (Seção V.2.1); (ii) comparar o desempenho do RTAD como método independente contra abordagens clássicas (Seção V.2.3); (iii) analisar a sensibilidade do método em relação à forma como a dispersão estatística é calculada (Seção V.2.4); e (iv) investigar os efeitos do critério de limiarização utilizado na etapa de decisão (Seção V.2.5).

V.2.1 Eficácia da RT

Esta etapa avalia o impacto da aplicação da RT a diferentes métodos clássicos de detecção de anomalias. O objetivo é determinar se a transformação melhora as capacidades de detecção dos métodos avaliados. Para isso, as métricas obtidas por cada método são comparadas na série original e na série transformada com RT. Os tamanhos de efeito também são calculados para quantificar a magnitude das melhorias promovidas por RT.

A Figura V.2 apresenta uma análise do impacto da RT sobre o desempenho dos métodos de detecção de anomalias no conjunto de séries temporais heteroscedásticas. Os subgráficos (a), (b) e (c) exibem a comparação para F1, revocação e precisão, respectivamente. Os gráficos de dispersão ilustram o desempenho de cada série temporal individual: o eixo horizontal representa o valor da métrica obtido sem a aplicação da RT (cenário "original"), enquanto o eixo vertical corresponde ao valor da métrica com a aplicação da RT. A linha tracejada vermelha, que representa a igualdade, serve como referencial; pontos situados acima dessa linha indicam que a aplicação da RT resultou em melhoria no desempenho para aquela série temporal específica, enquanto pontos abaixo sugerem deterioração. Pontos próximos à linha de igualdade denotam impacto neutro da RT. Adicionalmente, o 'X' vermelho em cada subgráfico indica as médias dos valores da métrica nos dois cenários, oferecendo uma referência visual imediata para o desempenho médio agregado. Essa abordagem fornece uma perspectiva granular e direta sobre a consistência e a magnitude das mudanças induzidas pela RT em nível de série temporal.

Observa-se que a RT demonstra potencial positivo e consistente na otimização das métricas para a maioria dos métodos avaliados. A revocação exibe um dos maiores benefícios, com aumento para quase todas as séries na maioria dos métodos após a aplicação da RT, indicando capacidade aprimorada dos modelos de detectar anomalias verdadeiras. A precisão também é influenciada

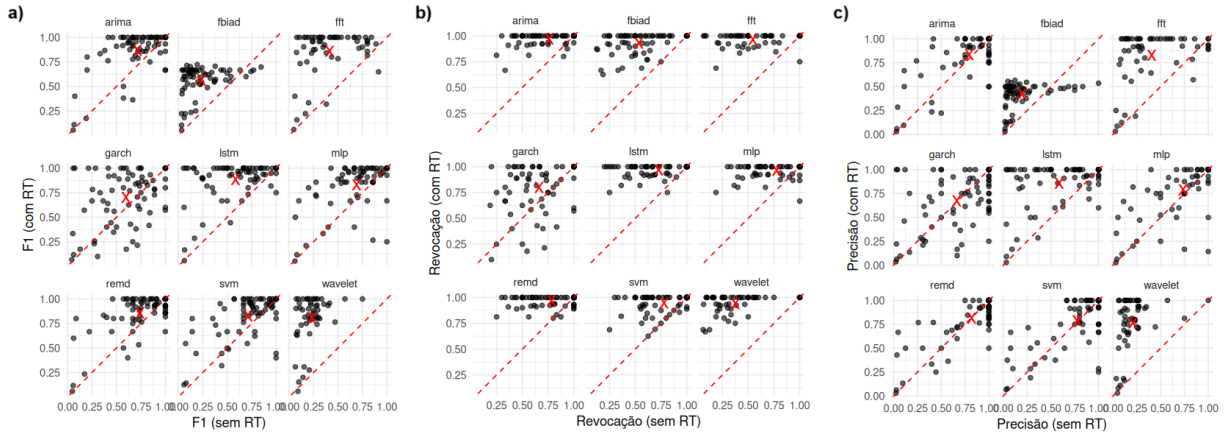


Figura V.2: Comparação do desempenho de diferentes modelos: (a) F1, (b) revocação e (c) precisão, com e sem aplicação da RT. Cada ponto representa uma série temporal; a linha tracejada indica igualdade de desempenho, e o 'X' indica a média dos resultados para cada modelo.

positivamente, embora com grau de melhoria mais variável entre os métodos. Consequentemente, o F1 apresenta resultados favoráveis à aplicação da RT, sugerindo melhor equilíbrio entre a detecção precisa de anomalias e a minimização de falsas detecções.

Entre os métodos avaliados, FBIAD, FFT, LSTM e WAVELET destacam-se como os principais beneficiários da RT, demonstrando melhorias substanciais nas três métricas para a maioria das séries temporais. WAVELET, em particular, evolui de desempenho inferior para níveis competitivos após a transformação. O FBIAD, embora apresente grande melhora, mantém desempenho relativamente inferior aos demais métodos. ARIMA, MLP, REMD e SVM, mesmo já exibindo alto desempenho sem RT, também apresentam melhoria geral, principalmente em revocação, muitas vezes atingindo o valor máximo. O GARCH, embora apresente ligeira melhora nas médias com RT, mostra ganhos menos significativos, com proporção considerável de séries sem alteração ou com perdas, principalmente em precisão.

De forma geral, observa-se predominância de pontos (e médias) acima da linha de igualdade, indicando que o desempenho *com* RT supera o desempenho *sem* a transformação na maioria dos métodos avaliados.

A Tabela V.2 resume os resultados da análise estatística comparando os métodos de detecção aplicados às séries originais e às versões transformadas pela RT. Para cada combinação de método e métrica (F1, precisão e revocação), o *status* e o tamanho do efeito correspondente são relatados. Conforme descrito na Seção V.1.5, o teste de Wilcoxon pareado é empregado com hipótese alternativa bilateral. A coluna *status* consolida os achados considerando significância estatística ($p \leq 0{,}05$) e a direção da mediana das diferenças entre desempenhos com e sem RT. Especificamente, "+" indica melhoria estatisticamente significativa; "−" denota deterioração significativa; "N" indica efeito estatisticamente significativo com mediana da diferença igual a zero (sem tendência

direcional clara, sugerindo efeitos heterogêneos entre séries); e "NS" implica ausência de evidência estatística de diferença. Tamanhos de efeito são categorizados como desprezível (d), pequeno (p), médio (m) ou grande (g).

Tabela V.2: Análise estatística comparativa de RT para métodos de detecção de anomalias em séries temporais heteroscedásticas.

	F1			**Precisão**			**Revocação**		
	status	**r**	**def**	**status**	**r**	**def**	**status**	**r**	**def**
ARIMA	+	0,61	g	NS	0,22	p	+	0,73	g
FBIAD	+	0,84	g	+	0,72	g	+	0,86	g
FFT	+	0,85	g	+	0,81	g	+	0,85	g
GARCH	+	0,30	p	NS	0,02	d	+	0,44	m
LSTM	+	0,80	g	+	0,68	g	+	0,76	g
MLP	+	0,58	g	N	0,32	m	+	0,67	g
REMD	+	0,53	g	NS	0,07	d	+	0,65	g
SVM	+	0,55	g	NS	0,15	p	+	0,65	g
WAVELET	+	0,87	g	+	0,87	g	+	0,87	g

Legenda: status = interpretação do teste de Wilcoxon pareado ($p \leq 0{,}05$) e direção da mediana das diferenças. +: melhoria; −: deterioração; **N**: efeito significativo sem direção (mediana zero); **NS**: não significativo. **r** = tamanho do efeito; **def**: **d** (desprezível, $r < 0{,}1$), **p** (pequeno), **m** (médio), **g** (grande).

A análise revelou um padrão consistente: sempre que uma diferença estatisticamente significativa foi identificada, as métricas de desempenho com RT foram superiores às das séries originais, indicando melhoria na detecção de anomalias proporcionada pela transformação. A maioria dos métodos mostra efeitos significativos com magnitudes ao menos médias, especialmente para F1 e revocação. Notavelmente, FBIAD, FFT e WAVELET exibem efeitos grandes em todas as métricas. ARIMA, LSTM, MLP, REMD e SVM também demonstram ganhos relevantes em pelo menos duas métricas. Apenas GARCH produz resultados mais limitados, com efeito pequeno em F1 e médio em revocação.

Em complemento às análises quantitativas e estatísticas, a Figura V.3 oferece uma ilustração prática do impacto da RT na detecção de anomalias em uma série temporal específica. Esta representação visual demonstra, em detalhe, como a aplicação da RT modifica os resultados de detecção para métodos selecionados.

A figura é composta por um painel superior que exibe a série temporal analisada. Abaixo deste, seis painéis de linha do tempo detalham os resultados da detecção para cada cenário, comparando o desempenho dos métodos FBIAD, ARIMA e GARCH. Essa representação qualitativa complementa as métricas quantitativas, oferecendo uma compreensão visual de como a RT refina o processo de detecção.

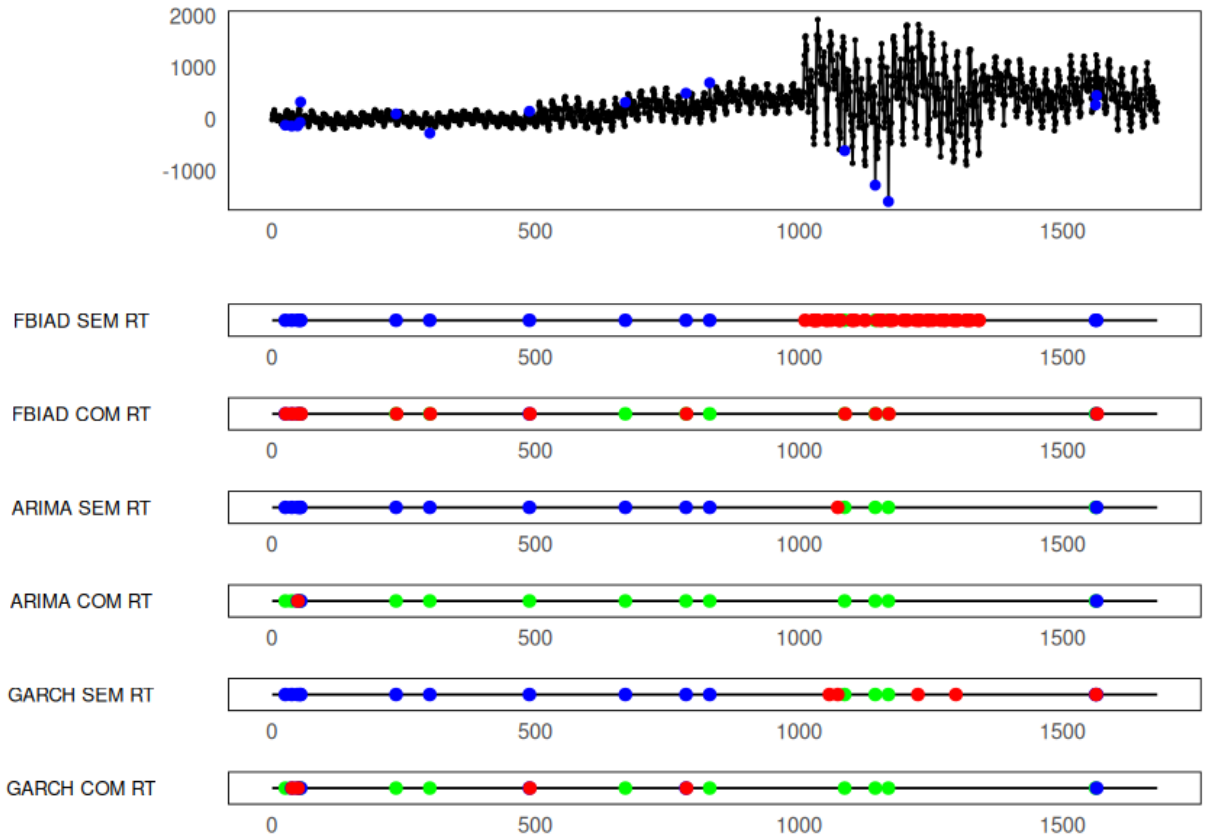


Figura V.3: Exemplo comparativo da detecção de anomalias com e sem aplicação da RT para os métodos FBIAD, ARIMA e GARCH em uma série temporal selecionada. O painel superior exibe a série temporal original. Os painéis inferiores detalham os resultados para cada configuração. Anomalias corretamente detectadas (Verdadeiros Positivos) são indicadas em verde, anomalias não detectadas (Falsos Negativos) em azul, e falsas detecções (Falsos Positivos) em vermelho.

V.2.2 Teste de Ablação: Análise no Conjunto de Séries Homoscedásticas

Para avaliar se a RT mantém sua utilidade ou, no mínimo, não prejudica o desempenho em cenários menos ideais, um teste foi conduzido no conjunto de séries homoscedásticas. Nesse contexto, esperava-se que os benefícios da transformação fossem menos pronunciados, servindo como um teste de ablação para a contribuição da RT sob condições em que a heteroscedasticidade não é fator predominante. Analogamente à Tabela V.2, a Tabela V.3 detalha os resultados da análise estatística para séries homoscedásticas, comparando métodos de detecção de anomalias com e sem a transformação RT.

Como previsto, na ausência de heteroscedasticidade, os benefícios da RT mostraram-se menos generalizados. Esse cenário serviu como teste de ablação, confirmando que a utilidade da RT é fortemente dependente do contexto. Contudo, mesmo em séries homoscedásticas, certos métodos — notadamente FFT, LSTM e WAVELET — demonstraram melhorias significativas em seu desempenho, com tamanhos de efeito grandes na maioria das métricas, o que sugere um impacto positivo da transformação que transcende a correção da heteroscedasticidade para esses modelos.

Tabela V.3: Análise estatística comparativa da RT para métodos de detecção de anomalias em séries temporais homoscedásticas.

	F1			Precisão			Revocação		
	status	**r**	**def**	**status**	**r**	**def**	**status**	**r**	**def**
ARIMA	N	0,31	m	N	0,31	m	NS	0,14	p
FBIAD	NS	0,02	d	−	0,25	p	+	0,67	g
FFT	+	0,82	g	+	0,79	g	+	0,86	g
GARCH	−	0,69	g	−	0,69	g	−	0,68	g
LSTM	+	0,69	g	+	0,70	g	N	0,52	g
MLP	N	0,25	p	N	0,25	p	NS	0,08	d
REMD	N	0,55	g	N	0,49	m	N	0,34	m
SVM	N	0,35	m	N	0,23	p	N	0,41	m
WAVELET	+	0,87	g	+	0,86	g	+	0,87	g

Legenda: status = interpretação do teste de Wilcoxon pareado ($p \leq 0{,}05$) e direção da mediana das diferenças. +: melhoria; −: deterioração; **N**: efeito significativo sem direção (mediana zero); **NS**: não significativo. **r** = tamanho do efeito; **def**: **d** (desprezível, $r < 0{,}1$), **p** (pequeno), **m** (médio), **g** (grande).

Em contraste, o método GARCH sofreu degradação consistente e estatisticamente significativa, com tamanhos de efeito grandes em todas as métricas. Para o FBIAD, a transformação resultou em degradação significativa na precisão, mas melhoria substancial na revocação, enquanto o F1 não apresentou diferença estatística significativa. ARIMA, MLP, REMD e SVM exibiram predominantemente diferenças estatisticamente significativas sem direção clara (status N) ou ausência de diferenças estatísticas relevantes.

Em síntese, o teste de ablação em séries homoscedásticas reitera que a RT não constitui uma solução universal de aprimoramento. Seus efeitos são específicos ao método, variando de benefícios acentuados (*e.g.*, FFT, WAVELET), passando por impactos mistos ou não direcionais, até prejuízos claros (*e.g.*, GARCH). Essa diversidade de resultados valida a premissa de que a principal vantagem da RT reside em cenários heteroscedásticos, mas também destaca sua capacidade de impactar o desempenho de modelos específicos mesmo quando essa condição não está presente, reafirmando sua utilidade direcionada e não indiscriminada.

V.2.3 Detector de Anomalias por Transformação Resiliente (RTAD)

Esta etapa avalia a capacidade da RT de servir como um método de detecção abrangente, sem depender de técnicas externas de detecção de anomalias. Para isso, um critério de decisão baseado em um modelo Gaussiano é aplicado diretamente à série transformada α_t, obtida a partir das séries temporais heteroscedásticas. A aplicação da RT, combinada com essa etapa de decisão, define um procedimento de detecção independente, referido como RTAD, cuja eficácia é comparada com métodos tradicionais aplicados às séries originais. A Figura V.4 mostra os resultados das métricas para cada método de detecção.

A análise da Figura V.4 demonstra que o método RTAD atinge a maior mediana do F1, resul-

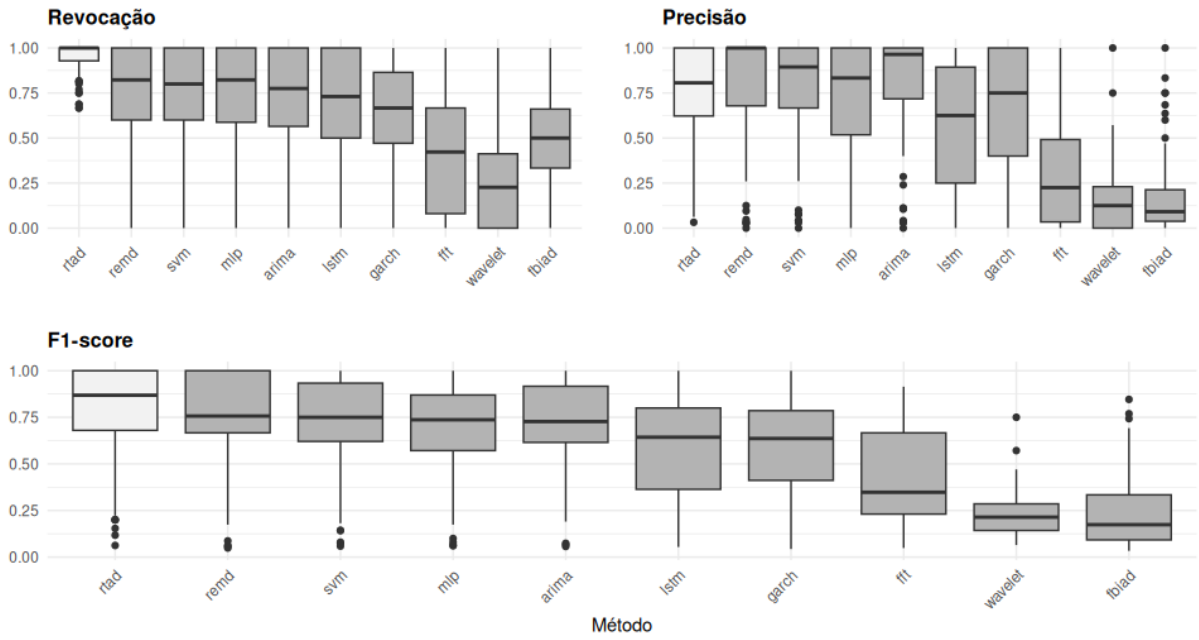


Figura V.4: Boxplots comparando F1, precisão e revocação de métodos de detecção para séries temporais heteroscedásticas. O método RTAD é destacado em cinza claro.

tado de uma combinação de precisão satisfatória e alta revocação. Em contraste, métodos como REMD, MLP, SVM e ARIMA, embora exibam medianas de precisão superiores às do RTAD, apresentam valores de revocação significativamente menores e mais dispersos. Essa distinção reflete uma estratégia de detecção mais conservadora, em que anomalias verdadeiras são frequentemente negligenciadas (falsos negativos). Tal revocação comprometida é uma limitação comum e desafiadora em séries temporais heteroscedásticas, devido à dificuldade em definir limiares que se adaptem simultaneamente a segmentos de alta e baixa variância, uma tarefa na qual métodos mais conservadores tendem a falhar.

Além disso, métodos de processamento de sinal, incluindo os algoritmos FFT e WAVELET, bem como o algoritmo FBIAD, exibem consistentemente as menores medianas em todas as métricas, indicando sua inadequação frente à complexidade da tarefa de detecção de anomalias neste contexto específico.

Assim como na Seção V.2.1, a análise do tamanho do efeito também é realizada. A Figura V.5 apresenta os tamanhos de efeito obtidos para cada combinação de método e métrica. Células sublinhadas indicam valores estatisticamente significativos ($p < 0{,}05$). Células com asterisco (*) indicam casos em que o valor mediano da métrica para RTAD é menor do que o do método comparado, destacando situações em que o RTAD teve desempenho inferior. Casos com sublinhado e asterisco indicam diferenças estatisticamente significativas desfavoráveis ao RTAD.

Na Figura V.5, os testes de significância estatística para as três métricas indicam que o RTAD é significativamente superior à maioria dos métodos. Exceções são encontradas apenas na métrica de

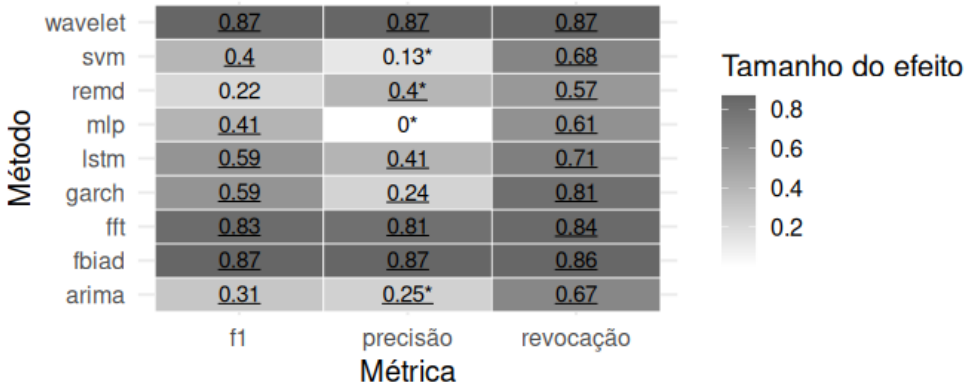


Figura V.5: Tamanho do efeito para cada método e métrica. Valores sublinhados indicam diferenças estatisticamente significativas; asteriscos indicam casos em que o valor mediano da métrica para RTAD é menor do que o do método correspondente.

precisão ao compará-lo com SVM, REMD, MLP e ARIMA. Os maiores efeitos são observados para WAVELET, FFT e FBIAD, com valores acima de 0,81, evidenciando a superioridade do RTAD. SVM, MLP, LSTM e GARCH mostram efeitos moderados. ARIMA e REMD exibem os menores tamanhos de efeito. A revocação é a métrica em que o RTAD mais se destaca.

Em resumo, quando combinada com um limiar estatístico, a RT pode servir como um método de detecção independente (RTAD) e supera os métodos tradicionais nas métricas consideradas para o conjunto de séries heteroscedásticas. Esses resultados reforçam o potencial do RTAD como uma abordagem autônoma para detecção de anomalias.

V.2.4 Sensibilidade do RTAD à Medida de Dispersão

Para avaliar a sensibilidade do RTAD, investigamos o impacto da medida de dispersão escolhida para construir a série transformada, a qual é a base da RT. Isso permite entender como diferentes quantificações da volatilidade afetam o desempenho final do detector. Dez abordagens para calcular a dispersão estatística móvel são consideradas: (i) desvio padrão, (ii) variância, (iii) desvio padrão da média móvel exponencialmente ponderada (EWMA), (iv) desvio absoluto mediano (MAD), (v) desvio absoluto médio (MeanAD), (vi) amplitude, (vii) amplitude interquartil (IQR), (viii) modelo baseado em GARCH, (ix) volatilidade realizada e (x) estimação baseada em wavelet.

A seleção dessas dez abordagens cobre um espectro diverso de medidas de dispersão, variando de estatísticas clássicas e robustas a modelos econométricos e técnicas de processamento de sinal. A Tabela V.4 resume as medidas de dispersão empregadas e suas principais contribuições para a análise de sensibilidade.

A Figura V.6 apresenta os valores de F1 obtidos pelo RTAD no conjunto de dados heteroscedástico, quando diferentes medidas de dispersão são utilizadas para o cálculo da volatilidade na etapa de RT. A Tabela V.5 reporta a porcentagem de séries heteroscedásticas que se tornam homoscedás-

Tabela V.4: Descrição e justificativa das medidas de dispersão consideradas.

Medida de Dispersão	**Justificativa/Contribuição**
Desvio Padrão (DP)	Clássica, linha de base; mesma unidade dos dados; interpretação direta da volatilidade.
Variância (VAR)	Clássica, linha de base; unidade ao quadrado; permite análise comparativa de escala.
Volatilidade EWMA	Dinâmica, ponderada temporalmente; enfatiza observações mais recentes.
Desvio Absoluto Mediano (MAD)	Robusta; menos sensível a *outliers*; adequada para dados com anomalias.
Desvio Absoluto Médio (MeanAD)	Robusta; alternativa ao MAD, baseada em desvios centrados na média.
Amplitude (AMP)	Simples e direta; captura a variação bruta máxima.
Amplitude Interquartil (IQR)	Robusta; foca na dispersão central, ignorando extremos.
Modelo GARCH	Econométrico; modela heteroscedasticidade e agrupamento de volatilidade.
Volatilidade Realizada (REA)	Direta; volatilidade não paramétrica a partir de dados de alta frequência.
Estimação Baseada em Wavelet	Multiescala; caracteriza a dispersão entre componentes de frequência.

ticas após a transformação, com base no teste estatístico de Breusch–Pagan (vide Seção II.1.4).

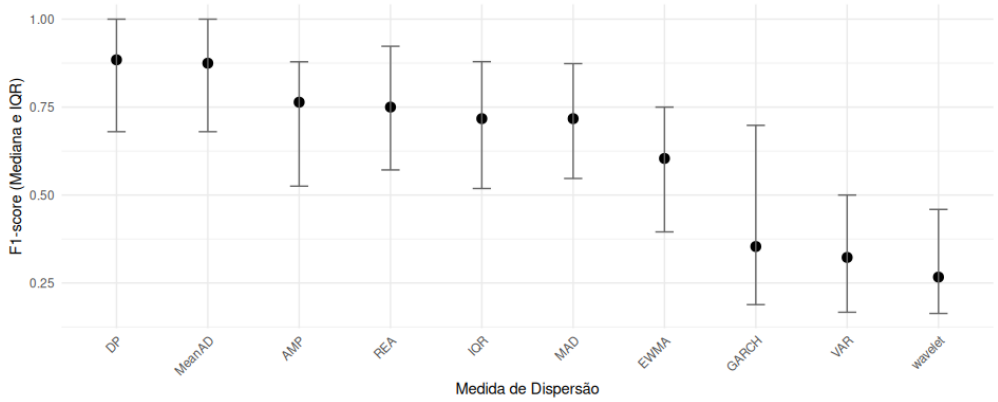


Figura V.6: F1 do RTAD no conjunto de dados heteroscedástico usando diferentes medidas de dispersão para cálculo da volatilidade.

Os resultados indicam que o desvio padrão proporciona o melhor desempenho de detecção (maior mediana de F1), seguido pelo desvio absoluto médio. Em contraste, medidas como IQR e amplitude, embora eficazes na normalização da variância (vide Tabela V.5), produzem valores de F1 mais baixos. Isso sugere que, apesar de reduzirem a heteroscedasticidade, essas medidas podem suavizar excessivamente a série, suprimindo anomalias em vez de destacá-las.

Tabela V.5: Redução percentual de séries heteroscedásticas por medida de dispersão, com base em 74 séries no conjunto de dados.

Medida de Dispersão	**Breusch–Pagan**
DP	95,9%
VAR	10,8%
EWMA	17,6%
MAD	93,2%
MeanAD	94,6%
AMP	97,3%
IQR	94,6%
GARCH	45,9%
REA	73,0%
wavelet	93,2%

Em resumo, a capacidade da RT de normalizar a variância e tornar uma série homoscedástica

não é, por si só, indicativa de melhor detecção de anomalias pelo RTAD. É necessário um equilíbrio entre a suavização da variância e a preservação do sinal de anomalias, que é o objetivo central da RT quando incorporada ao método RTAD.

V.2.5 Sensibilidade ao Critério de Limiar

Esta seção examina o impacto do critério de limiarização empregado na etapa de decisão do RTAD. Três estratégias são comparadas: limiarização baseada em um modelo Gaussiano, uma regra de *boxplot* e um critério de razão. Essas são as três técnicas de limiarização implementadas no Harbinger, com o modelo Gaussiano servindo como padrão. A Figura V.7 mostra os resultados para o conjunto de dados heteroscedástico.

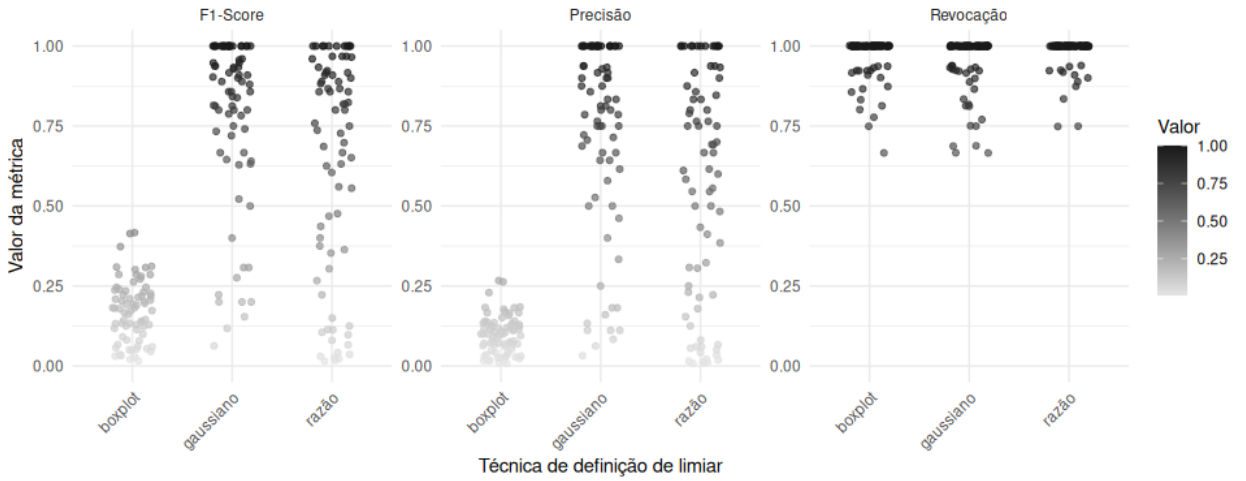


Figura V.7: Métricas de avaliação no conjunto de dados heteroscedástico utilizando diferentes estratégias de limiarização.

Os resultados mostram que o modelo Gaussiano proporciona o melhor desempenho geral, com alta mediana de F1 e precisão e baixa variabilidade. O critério baseado em razão produz desempenho semelhante, porém com menor estabilidade. O limiar baseado em *boxplot*, embora atinja alta revocação, resulta em F1 e precisão mais baixos, refletindo maior taxa de falsos positivos. Os resultados obtidos no conjunto de dados homoscedástico seguem o mesmo padrão.

A escolha da estratégia de limiarização não altera a conclusão da Seção V.2.3: o RTAD permanece superior aos demais métodos no conjunto de dados heteroscedástico, independentemente do critério de decisão aplicado (desde que aplicado de forma consistente a todos os métodos).

V.3 Discussão

A análise comparativa dos métodos aplicados a séries transformadas por RT, apresentada na Seção V.2.1, revela padrões explicáveis pelas características intrínsecas de cada abordagem. Os

métodos de melhor desempenho, notadamente WAVELET, FFT e FBIAD, operam no domínio da frequência ou extraem representações espectrais do sinal. Essas técnicas são intrinsecamente sensíveis a oscilações de alta frequência, o que as torna altamente compatíveis com a representação do sinal produzida pela RT. Em particular, o método WAVELET se beneficia do aprimoramento da separação de componentes proporcionado pela transformação, o que facilita sua operação e explica seu desempenho consistentemente elevado.

Em contraste, métodos como ARIMA, SVM, MLP e REMD exibem ganhos mais modestos. Métodos de aprendizado de máquina, por exemplo, são naturalmente capazes de se adaptar a padrões temporais complexos, o que lhes permite compensar parcialmente a ausência de pré-processamento. Contudo, essa flexibilidade adaptativa frequentemente implica um custo computacional elevado, limitando a escalabilidade. Ao produzir uma representação mais homogênea e com variância estabilizada, a RT pode, de fato, oferecer suporte a esses métodos. Ainda assim, a eficácia desse suporte depende da interação específica com o modelo de aprendizado, o que justifica a variabilidade nas melhorias observadas.

O caso do REMD merece atenção especial. Conceitualmente, REMD e RT compartilham vários elementos metodológicos, incluindo a decomposição por CEEMD, a análise baseada em rugosidade e a seleção de IMFs guiada pela curvatura. Devido a essa sobreposição, a aplicação da RT ao REMD pode resultar em operações redundantes, limitando o potencial de ganhos adicionais. Isso explica as melhorias relativamente menores observadas no REMD em comparação com outras abordagens de referência.

A interação entre GARCH e RT produz resultados distintos entre os tipos de séries. Para séries homoscedásticas, a RT degrada o desempenho do GARCH na detecção de anomalias. Esse comportamento é atribuído ao *design* inerente do GARCH, voltado para modelar a variância condicional. Em séries estáveis, a ausência de volatilidade estrutural limita sua capacidade preditiva, e a RT pode introduzir distorções desnecessárias. Por outro lado, para séries heteroscedásticas, a RT aprimora o F1 e a revocação do GARCH. Nesse contexto, a RT estabiliza a variância extrema, permitindo que o GARCH identifique anomalias de forma mais eficaz, apesar de suas abordagens conceitualmente distintas para o tratamento da volatilidade.

Os resultados da Seção V.2.3 demonstram que a RT alcança desempenho competitivo sem o auxílio de detectores externos, estabelecendo-se como uma solução autônoma de detecção de anomalias por meio da RTAD. Sua eficácia é comprovada em séries heteroscedásticas, destacando sua capacidade de normalizar a variabilidade local ao mesmo tempo em que preserva padrões relacionados a anomalias.

Os estudos de sensibilidade apresentados nas Seções V.2.4 e V.2.5 fornecem uma compreensão mais aprofundada das escolhas de *design* técnico que influenciam o desempenho tanto da RT quanto

do RTAD. Para as medidas de dispersão, os resultados indicam que o desvio padrão, amplamente utilizado na literatura, oferece o balanço mais favorável entre a estabilização da variância e a preservação de anomalias. Em relação às estratégias de limiarização, o modelo Gaussiano emerge como o mais eficaz, pois combina altos valores medianos de F1 com baixa variabilidade. Esse desempenho superior do modelo Gaussiano é atribuído ao alinhamento com o comportamento mais regularizado e homogêneo das séries após a RT. Embora a normalidade estrita da distribuição dos valores transformados não seja uma premissa direta da RT, a estabilização da variância e a homogeneização da amplitude promovidas pela transformação aproximam as séries de um estado de homoscedasticidade. Esta condição é particularmente favorável para a aplicação de um modelo de limiarização Gaussiano, que se baseia na estimativa de um desvio padrão mais estável para estabelecer seus limites, resultando na produção de limiares que minimizam tanto falsos positivos quanto falsos negativos.

Capítulo VI Conclusões

Este estudo aborda diretamente a questão de pesquisa apresentada na introdução: *"É possível aprimorar a detecção de anomalias em séries temporais heteroscedásticas por meio de uma transformação que equalize a variabilidade local?"*. Os achados fornecem uma resposta afirmativa. Demonstrou-se que tal aprimoramento é efetivamente alcançado por meio da *Resilient Transformation*. A transformação proposta integra a decomposição via CEEMD, a seleção automática de componentes de alta frequência, a diferenciação e a normalização pela dispersão local. A série resultante exibe amplitude aproximadamente homogênea ao longo do tempo, aumentando a visibilidade de anomalias subjacentes e cumprindo a premissa central de equalizar a variabilidade local.

A avaliação experimental confirma que a RT melhora de forma consistente o desempenho de diversos métodos de detecção de anomalias. Abordagens baseadas em resíduos e em representações espectrais, em particular, beneficiam-se do realce das oscilações de alta frequência. Esses ganhos destacam a RT como um passo de pré-processamento versátil, capaz de ser integrado a uma ampla gama de pipelines de detecção.

Adicionalmente, quando combinada com uma regra de decisão, a RT forma um método de detecção de anomalias totalmente autônomo, o RTAD. Os resultados demonstram que o RTAD supera significativamente os métodos tradicionais em séries temporais heteroscedásticas, alcançando o maior F1 mediano, desempenho impulsionado principalmente por sua alta revocação.

As análises de sensibilidade reforçam a solidez metodológica da RT e do RTAD, mostrando que o uso do desvio padrão como medida de dispersão e do limiar Gaussiano como critério de decisão produz resultados mais estáveis e eficazes. Esses achados confirmam que o *framework* proposto não é apenas conceitualmente sólido, mas também praticamente confiável.

Em suma, tanto a RT quanto o RTAD representam contribuições metodológicas significativas para a detecção de anomalias em séries temporais heteroscedásticas. Elas combinam alta eficácia, simplicidade computacional e independência de premissas paramétricas específicas, fornecendo uma solução abrangente e reprodutível para o problema central investigado nesta pesquisa.

Embora a RT demonstre vantagens substanciais, o método apresenta limitações que merecem consideração. Primeiramente, a validação metodológica baseou-se em subconjuntos de dados específicos. A dificuldade em encontrar séries temporais heterocedásticas com anomalias rotuladas é uma limitação inerente, o que restringe a abrangência da validação. Adicionalmente, a RT pode

enfrentar desafios em séries temporais com baixa variabilidade ou comportamento excessivamente estacionário, visto que a técnica de decomposição empregada (via CEEMD) pode não conseguir gerar IMFs em quantidade suficiente, comprometendo a eficácia da transformação. Além disso, o RTAD demonstra particular dificuldade em lidar com intervalos prolongados de anomalias (ou anomalias de sequência), pois tende a detectar predominantemente os pontos de início e fim da sequência anômala, negligenciando a captura de sua extensão completa. Outro desafio relevante reside na complexidade computacional, especialmente em séries temporais extensas ou de alta frequência, devido ao custo inerente à decomposição e à análise iterativa para a geração das IMFs. Esses aspectos podem restringir a aplicabilidade da RT em determinados contextos, tornando essencial a avaliação criteriosa de sua adequabilidade antes da implementação.

As limitações e os resultados promissores do presente estudo abrem caminho para diversas direções de pesquisa futura, visando tanto o aprimoramento metodológico da RT quanto a expansão de sua aplicabilidade. Uma linha de investigação promissora envolve a reavaliação dos critérios para o particionamento das IMFs em componentes de alta e baixa frequência. Embora a combinação de rugosidade e curvatura máxima tenha se mostrado eficaz, a natureza estocástica inerente ao CEEMD na geração das IMFs sugere a exploração de características alternativas para a seleção, como variância, curtose ou combinações de múltiplos atributos, bem como o desenvolvimento de novos critérios de divisão que transcendam a curvatura máxima, como a definição de limiares de variância ou energia para a inclusão de IMFs na análise de anomalias.

Adicionalmente, análises preliminares indicam que o RTAD pode se beneficiar de critérios de limiarização mais rigorosos do que o modelo Gaussiano adotado nesta tese. Esse aprimoramento é crucial para elevar a precisão, identificada como um dos pontos de menor desempenho do RTAD, sem comprometer a detecção de verdadeiros positivos. Espera-se que essa melhoria na precisão resulte em um aumento substancial do F1.

Outra vertente relevante para trabalhos futuros é a expansão da aplicabilidade da RT e do RTAD, incluindo a adaptação para detecção de anomalias em modo online, um tema de crescente relevância diante da complexidade computacional envolvida na geração das IMFs. O recálculo completo das IMFs a cada novo dado pode inviabilizar respostas rápidas; portanto, soluções potenciais incluem estratégias de segmentação da série temporal, janelas deslizantes com atualização parcial ou técnicas de alocação incremental de dados recentes em IMFs pré-existentes. Também há oportunidade para expandir a RT para a análise multivariada, possibilitando a detecção de anomalias que emergem de interações entre múltiplas variáveis, e não apenas de forma isolada.

O desenvolvimento da RT e do RTAD nesta tese reflete não apenas a resposta a uma lacuna na literatura, mas também o aprofundamento e a maturidade da pesquisa em detecção de anomalias em séries temporais. Essa consolidação metodológica foi impulsionada por trabalhos anteriores,

como o de Baroni et al. [2024], que investigou a importância da contextualização e da variabilidade local para a detecção precisa de anomalias em dados de saúde. Durante o doutorado, a autora também contribuiu para a área com o repositório unificado UniTED [Lima et al., 2025], o híbrido REMD [Souza et al., 2024] e a investigação de métricas de avaliação robustas [Escobar et al., 2021]. Essas aplicações, incluindo o estudo da subnotificação de COVID-19 [Paixão et al., 2021] e a análise de esquemas de *pump-and-dump* em criptomoedas [Moura et al., 2024], evidenciam a preocupação central desta pesquisa com a contextualização e a heteroscedasticidade em séries temporais.

Assim, esta tese avança sobre essa base consolidada, oferecendo uma proposta inovadora e resiliente para equalizar a variabilidade local e otimizar a detecção de anomalias em cenários heteroscedásticos. O arcabouço teórico e as evidências que sustentam essas contribuições são detalhados na seção de referências bibliográficas que se segue.

Referências

Alsan, H. F., Guler, A. K., Yildiz, E., Kilinc, S., Camlidere, B., and Arsan, T. Network Traffic Anomaly Detection Using Quantile Regression with Tolerance. In *2023 IEEE International Black Sea Conference on Communications and Networking, BlackSeaCom 2023*, pages 300 – 305, 2023. https://doi.org/10.1109/BlackSeaCom58138.2023.10299728

Antunes, F., O'Sullivan, A., Rodrigues, F., and Pereira, F. A review of heteroscedasticity treatment with gaussian processes and quantile regression meta-models. *Springer Geography*, pages 141 – 160, 2017. https://doi.org/10.1007/978-3-319-40902-3_9

Au Yeung, J. F. K., Wei, Z.-k., Chan, K. Y., Lau, H. Y. K., and Yiu, K.-F. C. Jump detection in financial time series using machine learning algorithms. *Soft Computing*, 24(3):1789 – 1801, 2020. https://doi.org/10.1007/s00500-019-04006-2

Audibert, J., Marti, S., Guyard, F., and Zuluaga, M. A. From Univariate to Multivariate Time Series Anomaly Detection with Non-Local Information. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13114 LNAI:186 – 194, 2021. https://doi.org/10.1007/978-3-030-91445-5_12

Balestrassi, P., Paiva, A., Souza, A., and Turrioni, J. A multivariate descriptor for change point detection in nonlinear time series. In *16th Power Systems Computation Conference, PSCC 2008*, 2008.

Baroni, L. R., Scoralick, L., Reis, A., Belloze, K. T., Pedroso, M. D. M., Alves, R. F. S., Boccolini, C. S., de Moraes Mello Boccolini, P. M., and Ogasawara, E. S. A contextual-compositional approach to discover associations between health determinants and health indicators for neonatal mortality rate monitoring in situations of anomalies. *PLOS ONE*, 19(12):e0310413, 2024. https://doi.org/10.1371/journal.pone.0310413

Bartlett, M. S. Properties of Sufficiency and Statistical Tests. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 160(901):268–282, 1937.

Bauer, D. F. Constructing confidence sets using rank statistics. *Journal of the American Statistical Association*, 67(339):687 – 690, 1972. https://doi.org/10.1080/01621459.1972.10481279

Bishop, C. M. *Pattern Recognition and Machine Learning*. Springer New York, 2016.

Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307 – 327, 1986. https://doi.org/10.1016/0304-4076(86)90063-1

Boniol, P. and Palpanas, T. Series2Graph: graph-based subsequence anomaly detection for time series. *Proc. VLDB Endow.*, 13(12):1821–1834, 2020. https://doi.org/10.14778/3407790.3407792

Box, G. E. P. and Cox, D. R. An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2):211–243, 1964. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., and Ljung, G. M. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.

Braei, M. and Wagner, S. Anomaly Detection in Univariate Time-series: A Survey on the State-of-the-Art, 2020.

Breusch, T. S. and Pagan, A. R. A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, 47(5):1287–1294, 1979. https://doi.org/10.2307/1911963

Brown, M. B. and Forsythe, A. B. Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346):364 – 367, 1974. https://doi.org/10.1080/01621459.1974.10482955

Buda, T. S., Caglayan, B., and Assem, H. DeepAD: A generic framework based on deep learning for time series anomaly detection. In *Pacific-Asia conference on knowledge discovery and data mining*, volume 10937 LNAI, pages 577 – 588. Springer, 2018. https://doi.org/10.1007/978-3-319-93034-3_46

Cavaliere, G. Unit root tests under time-varying variances. *Econometric Reviews*, 23(3):259 – 292, 2004. https://doi.org/10.1081/ETC-200028215

Chandola, V., Banerjee, A., and Kumar, V. Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), 2009. https://doi.org/10.1145/1541880.1541882

Charles, A. and Darné, O. The accuracy of asymmetric GARCH model estimation. *International Economics*, 157:179 – 202, 2019. https://doi.org/10.1016/j.inteco.2018.11.001

Chen, X., Qiu, Q., Li, C., and Xie, K. GraphAD: A Graph Neural Network for Entity-Wise Multivariate Time-Series Anomaly Detection. In *SIGIR 2022 - Proceedings of the 45th International*

ACM SIGIR Conference on Research and Development in Information Retrieval, pages 2297 – 2302, 2022a. https://doi.org/10.1145/3477495.3531848

Chen, Z., Chen, D., Zhang, X., Yuan, Z., and Cheng, X. Learning Graph Structures With Transformer for Multivariate Time-Series Anomaly Detection in IoT. *IEEE Internet of Things Journal*, 9(12):9179 – 9189, 2022b. https://doi.org/10.1109/JIOT.2021.3100509

Cheong, M.-S., Wu, M.-C., and Huang, S.-H. Interpretable Stock Anomaly Detection Based on Spatio-Temporal Relation Networks with Genetic Algorithm. *IEEE Access*, 9:68302 – 68319, 2021. https://doi.org/10.1109/ACCESS.2021.3077067

Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*. Routledge, 2013.

Cooley, J. W. and Tukey, J. W. An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19(90):249 – 259, 1965. https://doi.org/10.1090/S0025-5718-1965-0178586-1

Cox, D. R. Statistical significance tests. *British Journal of Clinical Pharmacology*, 14(3):325 – 331, 1982. https://doi.org/10.1111/j.1365-2125.1982.tb01987.x

Cox, D. R., Spjøtvoll, E., Johansen, S., van Zwet, W. R., Bithell, J. F., Barndorff-Nielsen, O., and Keuls, M. The Role of Significance Tests [with Discussion and Reply]. *Scandinavian Journal of Statistics*, 4(2):49–70, 1977.

Cryer, J. D. and Chan, K.-S. *Time Series Analysis: With Applications in R*. Springer Science & Business Media, 2008.

Dani, M.-C., Jollois, F.-X., Nadif, M., and Freixo, C. Adaptive threshold for anomaly detection using time series segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9491:82 – 89, 2015. https://doi.org/10.1007/978-3-319-26555-1_10

Dong, K., Wang, X., and Yu, H. Anomaly detection of rotary kiln sensor time series data based on deep learning. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 12803, 2023. https://doi.org/10.1117/12.3009271

Engle, R. F. Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4):987–1007, 1982. https://doi.org/10.2307/1912773

ENGLE, R. F. and NG, V. K. Measuring and Testing the Impact of News on Volatility. *The Journal of Finance*, 48(5):1749 – 1778, 1993. https://doi.org/10.1111/j.1540-6261.1993.tb05127.x

Escobar, L., Salles, R., Lima, J., Gea, C., Baroni, L., Ziviani, A., Pires, P., Delicato, F., Coutinho, R., Assis, L., and Ogasawara, E. Evaluating Temporal Bias in Time Series Event Detection Methods. *Journal of Information and Data Management*, 12(3), 2021. https://doi.org/10.5753/jidm.2021.1968

Esling, P. and Agon, C. Time-series data mining. *ACM Computing Surveys*, 45(1), 2012. https://doi.org/10.1145/2379776.2379788

Fagerland, M. W. T-tests, non-parametric tests, and large studiesa paradox of statistical practice? *BMC Medical Research Methodology*, 12(1):78, 2012. https://doi.org/10.1186/1471-2288-12-78

Francq, C. and Zakoian, J.-M. *GARCH Models: Structure, Statistical Inference and Financial Applications*. John Wiley & Sons, 2019.

Garcia, S., Luengo, J., and Herrera, F. *Data Preprocessing in Data Mining*. Springer, 2014.

Garg, S., Kaur, K., Batra, S., Kaddoum, G., Kumar, N., and Boukerche, A. A multi-stage anomaly detection scheme for augmenting the security in IoT-enabled applications. *Future Generation Computer Systems*, 104:105 – 118, 2020. https://doi.org/10.1016/j.future.2019.09.038

Geiger, A., Liu, D., Alnegheimish, S., Cuesta-Infante, A., and Veeramachaneni, K. TadGAN: Time Series Anomaly Detection Using Generative Adversarial Networks. In *Proceedings - 2020 IEEE International Conference on Big Data, Big Data 2020*, pages 33 – 43, 2020. https://doi.org/10.1109/BigData50022.2020.9378139

GLOSTEN, L. R., JAGANNATHAN, R., and RUNKLE, D. E. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *The Journal of Finance*, 48(5):1779 – 1801, 1993. https://doi.org/10.1111/j.1540-6261.1993.tb05128.x

Goldfeld, S. M. and Quandt, R. E. Some Tests for Homoscedasticity. *Journal of the American Statistical Association*, 60(310):539 – 547, 1965. https://doi.org/10.1080/01621459.1965.10480811

Gujarati, D. N. *Essentials of Econometrics*. SAGE, 2021.

Hamilton, J. D. *Time Series Analysis*. Princeton University Press, Princeton, New Jersey, USA, 2 edition, 1994.

Han, J., Pei, J., and Tong, H. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, Cambridge, MA, 4th edition edition, 2022.

Hassan, M., Hossny, M., Nahavandi, S., and Creighton, D. Heteroskedasticity variance index. In *Proceedings - 2012 14th International Conference on Modelling and Simulation, UKSim 2012*, pages 135 – 141, 2012. https://doi.org/10.1109/UKSim.2012.28

Hojat, M. and Xu, G. A visitor's guide to effect sizes–statistical significance versus practical (clinical) importance of research findings. *Advances in health sciences education*, 9(3):241–249, 2004.

Hong, C.-Y., Caesar, M., Duffield, N., and Wang, J. Tiresias: Online anomaly detection for hierarchical operational network data. In *Proceedings - International Conference on Distributed Computing Systems*, pages 173 – 182, 2012. https://doi.org/10.1109/ICDCS.2012.30

Horváth, L., Kokoszka, P., and Zhang, A. Monitoring constancy of variance in conditionally heteroskedastic time series. *Econometric Theory*, 22(3):373 – 402, 2006. https://doi.org/10.1017/S0266466606060191

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Snin, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H. The empirical mode decomposition and the Hubert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 454(1971):903 – 995, 1998. https://doi.org/10.1098/rspa.1998.0193

Hyndman, R. J. and Athanasopoulos, G. *Forecasting: principles and practice*. OTexts, 2018.

Hyndman, R. J. and Khandakar, Y. Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software*, 27(3):1 – 22, 2008. https://doi.org/10.18637/jss.v027.i03

Jiang, F., Zhu, Q., and Tian, T. An ensemble interval prediction model with change point detection and interval perturbation-based adjustment strategy: A case study of air quality. *Expert Systems with Applications*, 222, 2023. https://doi.org/10.1016/j.eswa.2023.119823

Jiang, M., Munawar, M. A., Reidemeister, T., and Ward, P. A. Heteroscedastic models to track relationships between management metrics. In *2009 IFIP/IEEE International Symposium on Integrated Network Management, IM 2009*, pages 375 – 381, 2009. https://doi.org/10.1109/INM.2009.5188838

Kanarachos, S., Christopoulos, S.-R. G., Chroneos, A., and Fitzpatrick, M. E. Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform. *Expert Systems with Applications*, 85:292 – 304, 2017. https://doi.org/10.1016/j.eswa.2017.04.028

Kieu, T., Yang, B., Guo, C., Cirstea, R.-G., Zhao, Y., Song, Y., and Jensen, C. S. Anomaly Detection in Time Series with Robust Variational Quasi-Recurrent Autoencoders. In *Proceedings*

- *International Conference on Data Engineering*, volume 2022-May, pages 1342 – 1354, 2022. https://doi.org/10.1109/ICDE53745.2022.00105

Kim, D., Park, S., and Choo, J. When model meets new normals: test-time adaptation for unsupervised time-series anomaly detection. In *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence and Thirty-Sixth Conference on Innovative Applications of Artificial Intelligence and Fourteenth Symposium on Educational Advances in Artificial Intelligence*, volume 38 of *AAAI'24/IAAI'24/EAAI'24*, pages 13113–13121. AAAI Press, 2024. https://doi.org/10.1609/aaai.v38i12.29210

Kim, H. and Kim, H. Contextual anomaly detection for multivariate time series data. *Quality Engineering*, 35(4):686 – 695, 2023. https://doi.org/10.1080/08982112.2023.2179404

Lampert, C. H., Blaschko, M. B., and Hofmann, T. Beyond sliding windows: Object localization by efficient subwindow search. In *26th IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2008. https://doi.org/10.1109/CVPR.2008.4587586

Lima, J., Castro, H., Oliveira, L., Paixão, E., Baroni, L., Salles, R., Vargas, R., and Ogasawara, E. UniTED: A Unified Time Series Event Detection Repository. In *Brazilian e-Science Workshop (BreSci)*, pages 1–8. SBC, 2025. https://doi.org/10.5753/bresci.2025.247972

Lima, J., Salles, R., Porto, F., Coutinho, R., Alpis, P., Escobar, L., Pacitti, E., and Ogasawara, E. Forward and Backward Inertial Anomaly Detector: A Novel Time Series Event Detection Method. In *Proceedings of the International Joint Conference on Neural Networks*, volume 2022-July, 2022. https://doi.org/10.1109/IJCNN55064.2022.9892088

Lin, S., Clark, R., Birke, R., Schonborn, S., Trigoni, N., and Roberts, S. Anomaly Detection for Time Series Using VAE-LSTM Hybrid Model. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2020-May, pages 4322 – 4326, 2020.

Loperfido, N. Kurtosis maximization for outlier detection in GARCH Models. *Mathematical and Statistical Methods for Actuarial Sciences and Finance, MAF 2018*, pages 455 – 459, 2018. https://doi.org/10.1007/978-3-319-89824-7_81

Mari, C. and Baldassari, C. Ensemble methods for jump-diffusion models of power prices. *Energies*, 14(8), 2021. https://doi.org/10.3390/en14082084

McArdle, B. H. and Anderson, M. J. Variance heterogeneity, transformations, and models of species abundance: A cautionary tale. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(7):1294 – 1302, 2004. https://doi.org/10.1139/F04-051

Moore, D. S., McCabe, G. P., and Craig, B. A. *Introduction to the Practice of Statistics*. W. H. Freeman, 2014.

Moura, M. S., Baroni, L., Ogasawara, E., and Mendonça, D. S. HD Pump: A Hybrid Detection Approach for Pump-and-Dump Schemes in Cryptocurrency Exchanges. In *Simpósio Brasileiro de Banco de Dados (SBBD)*, pages 757–763. SBC, 2024. https://doi.org/10.5753/sbbd.2024.243293

Nelson, D. B. Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, 59(2):347–370, 1991. https://doi.org/10.2307/2938260

Nirmalkar, K. Wilcoxon Test in R: The Ultimate Guide, 2025.

Ogasawara, E., Castro, A., Mello, A., Paixão, E., Fraga, F., Borges, H., Lima, J., Souza, J., Baroni, L., Tavares, L., Reis, M., Salles, R., Carvalho, D., Bezerra, E., Coutinho, R., Pacitti, E., and Porto, F. harbinger: A Unified Time Series Event Detection Framework, 2025a.

Ogasawara, E., Martinez, L. C., De Oliveira, D., Zimbrão, G., Pappa, G. L., and Mattoso, M. Adaptive Normalization: A novel data normalization approach for non-stationary time series. In *Proceedings of the International Joint Conference on Neural Networks*, 2010. https://doi.org/10.1109/IJCNN.2010.5596746

Ogasawara, E., Salles, R., Porto, F., and Pacitti, E. *Event Detection in Time Series*. Synthesis Lectures on Data Management. Springer Nature Switzerland, Cham, 1 edition, 2025b.

Ou, T., Cao, J., Lu, Y., Wang, Y., and Wu, X. A GARCH-Based Correction Method for Supervised Time Series Anomaly Detection. In *2023 3rd International Conference on Consumer Electronics and Computer Engineering, ICCECE 2023*, pages 86 – 91, 2023. https://doi.org/10.1109/ICCECE58074.2023.10135404

Paixão, B., Baroni, L., Pedroso, M., Salles, R., Escobar, L., de Sousa, C., de Freitas Saldanha, R., Soares, J., Coutinho, R., Porto, F., and Ogasawara, E. Estimation of COVID-19 Under-Reporting in the Brazilian States Through SARI. *New Generation Computing*, 39(3-4):623 – 645, 2021. https://doi.org/10.1007/s00354-021-00125-3

Percival, D. B. and Walden, A. T. *Wavelet Methods for Time Series Analysis*. Cambridge University Press, Cambridge, 2000.

Pour, M. Determining surface roughness of machining process types using a hybrid algorithm based on time series analysis and wavelet transform. *International Journal of Advanced Manufacturing Technology*, 97(5-8):2603 – 2619, 2018. https://doi.org/10.1007/s00170-018-2070-2

Qiu, C., Pfrommer, T., Kloft, M., Mandt, S., and Rudolph, M. Neural Transformation Learning for Deep Anomaly Detection Beyond Images. In *Proceedings of the 38th International Conference on Machine Learning*, pages 8703–8714. PMLR, 2021.

Riani, M. Robust transformations in univariate and multivariate time Series. *Econometric Reviews*, 28(1-3):262 – 278, 2009. https://doi.org/10.1080/07474930802388074

Ridley, D. A model-free power transformation to homoscedasticity. *International Journal of Production Economics*, 36(2):191 – 202, 1994. https://doi.org/10.1016/0925-5273(94)90024-8

Rosenthal, R. Parametric measures of effect size. In *The handbook of research synthesis*, volume 621, pages 231–244. Russell Sage Foundation, New York, NY, US, 1994.

Rosenthal, R. and Rubin, D. B. r equivalent: A Simple Effect Size Indicator. *Psychological Methods*, 8(4):492 – 496, 2003. https://doi.org/10.1037/1082-989X.8.4.492

Saito, T. and Rehmsmeier, M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE*, 10(3), 2015. https://doi.org/10.1371/journal.pone.0118432

Salles, R., Belloze, K., Porto, F., Gonzalez, P. H., and Ogasawara, E. Nonstationary time series transformation methods: An experimental review. *Knowledge-Based Systems*, 164:274 – 291, 2019. https://doi.org/10.1016/j.knosys.2018.10.041

Salles, R., Escobar, L., Baroni, L., Zorrilla, R., Ziviani, A., Kreischer, V., Delicato, F., Pires, P. F., Maia, L., Coutinho, R., Assis, L., and Ogasawara, E. Harbinger: Um framework para integração e análise de métodos de detecção de eventos em séries temporais. In *Anais do Simpósio Brasileiro de Banco de Dados (SBBD)*, pages 73–84. SBC, 2020. https://doi.org/10.5753/sbbd.2020.13626

Sarkar, M. Ruggedness measures of medical time series using fuzzy-rough sets and fractals. *Pattern Recognition Letters*, 27(5):447 – 454, 2006. https://doi.org/10.1016/j.patrec.2005.09.007

Sentana, E. Quadratic arch models. *Review of Economic Studies*, 62(4):639 – 661, 1995. https://doi.org/10.2307/2298081

Shumway, R. H. and Stoffer, D. S. *Time Series Analysis and Its Applications: With R Examples*. Springer, 2017.

Smola, A. J. and Schölkopf, B. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199 – 222, 2004. https://doi.org/10.1023/B:STCO.0000035301.49549.88

Song, J.-h., Prado, M. L. D., Simon, H. D., and Wu, K. J. Exploring Irregular Time Series through Non-Uniform Fast Fourier Transform. In *2014 Seventh Workshop on High Performance Computational Finance*, pages 37 – 44. IEEE, 2014. https://doi.org/10.1109/WHPCF.2014.8

Souza, J., Paixão, E., Fraga, F., Baroni, L., Alves, R. F. S., Belloze, K., Dos Santos, J., Bezerra, E., Porto, F., and Ogasawara, E. REMD: A Novel Hybrid Anomaly Detection Method Based on EMD and ARIMA. In *Proceedings of the International Joint Conference on Neural Networks*, 2024. https://doi.org/10.1109/IJCNN60899.2024.10651192

Spokoiny, V. Multiscale local change point detection with applications to value-at-risk. *Annals of Statistics*, 37(3):1405 – 1436, 2009. https://doi.org/10.1214/08-AOS612

Sullivan, G. M. and Feinn, R. Using Effect Size—or Why the P Value Is Not Enough. *Journal of Graduate Medical Education*, 4(3):279–282, 2012. https://doi.org/10.4300/JGME-D-12-00156.1

Sun, B., Ma, L., Shen, T., Geng, R., Zhou, Y., and Tian, Y. A robust data-driven method for multiseasonality and heteroscedasticity in time series preprocessing. *Wireless Communications and Mobile Computing*, 2021, 2021. https://doi.org/10.1155/2021/6692390

Thill, M., Konen, W., and Back, T. Online anomaly detection on the webscope S5 dataset: A comparative study. In *IEEE Conference on Evolving and Adaptive Intelligent Systems*, volume 2017-May, 2017. https://doi.org/10.1109/EAIS.2017.7954844

Tomczak, E. and Tomczak, M. The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences*, 21(1), 2014.

Tuli, S., Casale, G., and Jennings, N. R. TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data. In *Contemporary Mathematics*, volume 15, pages 1201 – 1214, 2022. https://doi.org/10.14778/3514061.3514067

Wang, D., Fortier, P. J., Michel, H. E., and Mitsa, T. Hierarchical agglomerative clustering based t-outlier detection. In *Proceedings - IEEE International Conference on Data Mining, ICDM*, pages 731 – 735, 2006. https://doi.org/10.1109/icdmw.2006.91

Wang, H. and Zhang, H. An Anomaly Detection Method for Multivariate Time Series Data Based on Variational Autoencoders and Association Discrepancy. *Mathematics*, 13(7), 2025. https://doi.org/10.3390/math13071209

Wang, M., Lyu, B., and Yu, G. ConvexVST: A Convex Optimization Approach to Variance-stabilizing Transformation. In *Proceedings of Machine Learning Research*, volume 139, pages 10839 – 10848, 2021.

Wenig, P., Schmidl, S., and Papenbrock, T. Anomaly Detectors for Multivariate Time Series: The Proof of the Pudding is in the Eating. In *Proceedings - 2024 IEEE 40th International Conference on Data Engineering Workshops, ICDEW 2024*, pages 96 – 101, 2024. https://doi.org/10.1109/ICDEW61823.2024.00018

White, H. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4):817–838, 1980. https://doi.org/10.2307/1912934

Wooldridge, J. M. *Introductory Econometrics: A Modern Approach*. South-Western Cengage Learning, 2013.

Wu, Z. and Huang, N. E. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1):1 – 41, 2009. https://doi.org/10.1142/S1793536909000047

Xu, L., Xu, K., Qin, Y., Li, Y., Huang, X., Lin, Z., Ye, N., and Ji, X. TGAN-AD: Transformer-Based GAN for Anomaly Detection of Time Series Data. *Applied Sciences (Switzerland)*, 12, 2022. https://doi.org/10.3390/app12168085

Yeh, J.-R., Shieh, J.-S., and Huang, N. E. Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method. *Advances in Adaptive Data Analysis*, 2(2):135 – 156, 2010. https://doi.org/10.1142/S1793536910000422

Yin, C., Zhang, S., Wang, J., and Xiong, N. N. Anomaly Detection Based on Convolutional Recurrent Autoencoder for IoT Time Series. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(1):112–122, 2022. https://doi.org/10.1109/TSMC.2020.2968516

Zamanzadeh Darban, Z., Webb, G. I., Pan, S., Aggarwal, C., and Salehi, M. Deep Learning for Time Series Anomaly Detection: A Survey. *ACM Computing Surveys*, 57(1), 2024. https://doi.org/10.1145/3691338

Zeiler, A., Faltermeier, R., Keck, I., Tomé, A., Puntonet, C., and Lang, E. Empirical mode decomposition - An introduction. In *Proceedings of the International Joint Conference on Neural Networks*, 2010. https://doi.org/10.1109/IJCNN.2010.5596829

Zhang, C., Xie, B., and Huo, Z. Unsupervised Anomaly Detection in Time Series Data via Enhanced VAE-Transformer Framework. *Computers, Materials and Continua*, 84(1):843 – 860, 2025. https://doi.org/10.32604/cmc.2025.063151

Zhang, Y., Li, B., and Zhang, X. Deep Learning-Based Anomaly Detection for Time-Series Data in Industrial Control Systems. In *Chinese Control Conference, CCC*, volume 2023-July, pages 8825 – 8829, 2023. https://doi.org/10.23919/CCC58697.2023.10241095

Zhao, Q., Zhang, Y., Shi, Y., and Li, J. Analyzing and Visualizing Anomalies and Events in Time Series of Network Traffic. *Advances in Intelligent Systems and Computing*, 936:15 – 25, 2020. https://doi.org/10.1007/978-3-030-19861-9_2

Zou, B., Yang, K., Kui, X., Liu, J., Liao, S., and Zhao, W. Anomaly detection for streaming data based on grid-clustering and Gaussian distribution. *Information Sciences*, 638, 2023. https://doi.org/10.1016/j.ins.2023.118989