

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA  
DEPARTAMENTO DE ESTATÍSTICA

Análise de séries temporais multivariadas via  
*Wavelet*

Matheus Kengi Mori

Trabalho de Conclusão de Curso



FEDERAL UNIVERSITY OF SÃO CARLOS  
EXACT AND TECHNOLOGY SCIENCES CENTER  
DEPARTMENT OF STATISTICS

# Wavelet Multivariate Time Series Analysis

**Matheus Kengi Mori**

Advisor: Prof.<sup>a</sup> Dra. Maria Sílvia de Assis Moura

Bachelors dissertation submitted to the Department of Statistics, Federal University of São Carlos - DEs-UFSCar, in partial fulfillment of the requirements for the degree of Bachelor in Statistics.

São Carlos  
September 2023



Matheus Kengi Mori

## Multivariate time series analysis via *Wavelet*

Este exemplar corresponde à redação final do trabalho de conclusão de curso devidamente corrigido e defendido por nome do aluno e aprovado pela banca examinadora.

Aprovado em 24 de Agosto de 2023

Banca Examinadora:

- Profa. Dra. Maria Sílvia de Assis Moura – Orientadora
- Prof. Dr. Francisco Antonio Rojas Rojas – Membro da banca 1
- Prof. Dr. Luis Aparecido Milan – Membro da banca 2



*A todos aqueles que de alguma forma estiveram e estão próximos de mim, fazendo esta vida valer cada vez mais a pena.*





# Agradecimentos

Agradeço a minha professora orientadora que teve paciência e que me ajudou bastante a concluir este trabalho, agradeço também aos meus professores que durante muito tempo me ensinaram e que me mostraram o quanto estudar é bom. Além disso, agradeço aos meus pais e irmãos, por todo o apoio e pela ajuda, que muito contribuíram para a realização deste trabalho.



*Só aquele que está disposto a ir longe demais sabe o quão longe é possível ir.*

(Johann Faigle)



# Resumo

Depois de meados da década de 1980, a teoria *Wavelet* rapidamente se espalhou para vários campos. Embora a análise *Wavelet* tenha a capacidade de decompor os dados em várias escalas de tempo e de lidar com dados não estacionários e localização no tempo, essa metodologia não se tornou uma aplicação popular em economia.

O presente trabalho de conclusão de curso irá utilizar as metodologias de transformação *wavelet*, decomposição *wavelet* e coerência *wavelet* para analisar a correlação entre quatro séries financeiras, como o *Dow Jones*, S&P500, IBOVESPA e Bitcoin. A fim de incrementar as análises econômicas, uma técnica pouco utilizada no campo da economia.

**Palavras-chave:** *estatística multivariada, séries temporais, Wavelet, mapa de coerência, séries financeiras, correlação.*



# Abstract

After the mid-1980s, wavelet theory quickly spread to many fields. Although wavelet analysis has the ability to decompose data into various time scales and to deal with non-stationary data and time location, this methodology has not become a popular application in economics.

This course completion work will use wavelet transformation, wavelet decomposition and wavelet coherence methodologies to analyze the correlation between four financial series, such as the Dow Jones, S&P500, IBOVESPA and Bitcoin. In order to enhance economic analysis, a technique that is little used in the field of economics.

**Keywords:** *multivariate statistics, time series, Wavelets, coherence map, financial series, correlation .*





# Lista de Figuras

2.1	À esquerda a série original e à direita a série diferenciada . . . . .	26
2.2	A série temporal $Z(t)$ à esquerda e a série estabilizada $\log(Z(t))$ à direita. . .	27
2.3	À esquerda a <i>wavelet</i> de Gauss e à direita a <i>wavelet</i> de Haar. . . . .	28
2.4	Exemplos de funções <i>wavelet</i> mãe usadas na TCW. . . . .	30
2.5	Esquema de um plano tempo $\times$ frequência. . . . .	31
2.6	Representação da transformada <i>wavelet</i> de um sinal; (a) Série da frequência da aceleração; (b) visualização 3D. . . . .	32
3.1	Representação da transformada <i>wavelet</i> de um sinal; (a) visualização 3D; (b) visualização pelo escalograma . . . . .	37
3.2	Gráfico de duas ondas senoidais (10 Hz e 50 Hz) em ruído branco . . . . .	38
3.3	Gráfico do mapa de coerência entre dois sinais . . . . .	39
4.1	Gráfico da matriz de dados faltantes . . . . .	43
4.2	Gráfico do preço de fechamento do índice IBOVESPA de 2012 até 2023 . . .	44
4.3	Gráfico da média e variância móvel do preço de fechamento do índice IBOVESPA de 2012 até 2023 . . . . .	44
4.4	Gráfico do preço de fechamento do <b>Dow Jones</b> de 2012 até 2023 . . . . .	45
4.5	Gráfico da média e variância móvel do preço de fechamento do <i>Dow Jones</i> de 2012 até 2023 . . . . .	46
4.6	Gráfico do preço de fechamento do Bitcoin de 2012 até 2023 . . . . .	46
4.7	Gráfico da média e variância móvel do preço de fechamento do Bitcoin de 2012 até 2023 . . . . .	47
4.8	Gráfico do preço de fechamento do <b>Dow Jones</b> de 2012 até 2023 . . . . .	47
4.9	Gráfico da média e variância móvel do preço de fechamento do <i>Dow Jones</i> de 2012 até 2023 . . . . .	48

4.10	Gráfico dos preços de fechamento do <i>Dow Jones</i> e IBOVESPA de 2012 até 2023 . . . . .	48
4.11	Representação da transformada <i>wavelet</i> dos preços de fechamento do IBOVESPA de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma . . . . .	49
4.12	Representação da transformada <i>wavelet</i> dos preços de fechamento do <i>Dow Jones</i> de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma . . . . .	49
4.13	Representação da transformada <i>wavelet</i> dos preços de fechamento do Bitcoin de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma . . . . .	50
4.14	Representação da transformada <i>wavelet</i> dos preços de fechamento do S&P500 de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma . . . . .	50
4.15	Gráfico da <b>Wavelet</b> ARM do preço de fechamento do <i>Dow Jones</i> de 2012 até 2023 . . . . .	51
4.16	Gráfico da <b>Wavelet</b> ARM do preço de fechamento do IBOVESPA de 2012 até 2023 . . . . .	51
4.17	Gráfico da <b>Wavelet</b> ARM do preço de fechamento do S&P500 de 2012 até 2023 . . . . .	52
4.18	Gráfico da <b>Wavelet</b> ARM do preço de fechamento do S&P500 de 2012 até 2023 . . . . .	52
4.19	Gráfico da correlação multiescala dos preços de fechamento do IBOVESPA e <i>Dow Jones</i> de 2012 até 2023 . . . . .	53
4.20	Gráfico da correlação multiescala dos preços de fechamento do S&P500 e <i>Dow Jones</i> de 2012 até 2023 . . . . .	54
4.21	Gráfico da correlação multiescala dos preços de fechamento do Bitcoin e <i>Dow Jones</i> de 2012 até 2023 . . . . .	55
4.22	Gráfico do mapa de coerência dos preços de fechamento entre o IBOVESPA e <i>Dow Jones</i> de 2012 até 2023 . . . . .	56
4.23	Gráfico do mapa de coerência dos preços de fechamento entre o S&P500 e <i>Dow Jones</i> de 2012 até 2023 . . . . .	57
4.24	Gráfico do mapa de coerência dos preços de fechamento entre o Bitcoin e <i>Dow Jones</i> de 2012 até 2023 . . . . .	57

# Lista de Tabelas

4.1	Tabela da correlação multiescala dos preços de fechamento do IBOVESPA e <i>Dow Jones</i> de 2012 até 2023 . . . . .	53
4.2	Tabela da correlação multiescala dos preços de fechamento do S&P500 e <i>Dow Jones</i> de 2012 até 2023 . . . . .	54
4.3	Tabela da correlação multiescala dos preços de fechamento do Bitcoin e <i>Dow Jones</i> de 2012 até 2023 . . . . .	55



# Notações e convenções

## Abreviações

FAC função de autocorrelação

FACP função de autocorrelação parcial

FDE função de densidade espectral

TWC transformação de wavelets contínuas

TWD transformação de wavelets discretas

## Notações não gregas

$L^2$  Conjunto de funções quadradas integráveis

$N$  Tamanho amostral

$\mathbb{R}$  Conjunto dos números reais

$t$  Tempo real (contínuo) ou um índice sem unidade (discreto)

$\mathbb{Z}$  Conjunto dos números inteiros

## Notações gregas

$\epsilon$  Um número pequeno e positivo

$\psi(\cdot)$  função *wavelet*

$\phi(\cdot)$  função de escala

## Outras convenções matemáticas e símbolos

$\approx$  aproximadamente igual

$\in, \notin$  pertence a, não pertence a

$\langle \cdot, \cdot \rangle$  produto interno

$|z|$  modulo (valor absoluto) de  $z$

$[a, b]$  intervalo de valores de  $x$  tal que  $a \leq x \leq b$

$S_j(\cdot)$  FDE para  $\bar{w}_{j,t}$  ou para parte não delimitada de  $w_{j,t}$

$S_X(\cdot)$  função de densidade espectral (de potência)



# Sumário

<b>1</b>	<b>Introdução</b>	<b>23</b>
<b>2</b>	<b><i>Wavelets</i> e transformações</b>	<b>25</b>
2.1	Séries temporais . . . . .	25
2.2	<i>Wavelets</i> . . . . .	27
2.2.1	Funções <i>wavelets</i> . . . . .	28
2.2.2	Transformação <i>Wavelet</i> Contínua (TWC) . . . . .	30
2.2.3	Transformação <i>Wavelet</i> Discreta (TWD) . . . . .	32
2.3	Como escolher uma função <i>wavelet</i> . . . . .	33
<b>3</b>	<b>Correlação entre séries temporais via <i>wavelet</i></b>	<b>35</b>
3.1	Mapa de coerência de <i>wavelet</i> . . . . .	37
<b>4</b>	<b>Aplicação</b>	<b>41</b>
4.1	Banco de dados . . . . .	41
4.2	Procedimento metodológicos . . . . .	42
4.2.1	Análise descritiva . . . . .	43
4.2.2	Correlação . . . . .	51
4.2.3	Mapa de coerência . . . . .	55
<b>5</b>	<b>Conclusão</b>	<b>59</b>
	<b>Referências Bibliográficas</b>	<b>60</b>
<b>A</b>	<b>Apêndice</b>	<b>63</b>





# Capítulo 1

## Introdução

Pela grade curricular do curso de estatística da UFSCar os alunos se deparam por métodos que analisam séries temporais de maneira univariada. Existem métodos que possibilitam uma análise multivariada de séries temporais, além da possibilidade de se trabalhar com dados que são não estacionários.

Um destes métodos é da teoria de *wavelet*, que depois de meados da década de 1980 rapidamente se espalhou para várias áreas de conhecimento, como na física, epidemiologia, engenharia elétrica e matemática. Entretanto essa metodologia não se tornou uma aplicação popular em economia, segundo [Aguiar-Conraria \(2011\)](#).

Visto o poder deste método e o desejo de popularizar, então este estudo tem como foco aplicar transformação de *wavelet* para estudar as relações entre as séries financeiras dispostas no mercado brasileiro.

Este trabalho contem na sua primeira parte uma revisão bibliográfica sobre *wavelet* múltipla e coerência *wavelet* dupla. Na segunda parte, se descreve a análise de correlação entre algumas séries financeiras, como IBOVESPA, cotação do dólar, taxa de juros e IPCA.

Além disso, este estudo utiliza ferramentas da metodologia *wavelet* que consiste em transformação contínua de *wavelet*, decomposição de séries via *wavelet*, correlação multi-escala, coerência de *wavelet*.

Esse trabalho está assim estruturado: no capítulo 2, temos a definição de séries temporais, apresentamos conceitos básicos sobre *wavelet*, transformações via *wavelet* e correlação via *wavelet*. No capítulo 3 fazemos a aplicação para as séries financeiras BOVESPA, *Dow Jones*, Bitcoin e S&P500. E uma conclusão do trabalho é apresentada no capítulo 4.



# Capítulo 2

## *Wavelets* e transformações

### 2.1 Séries temporais

Antes de entrar em mais detalhes sobre *wavelets*, é importante que o leitor entenda os principais conceitos de análise de séries temporais. Logo, segundo MORETTIN (2018) série temporal é qualquer conjunto de observações ordenada no tempo. E existem várias séries temporais que são utilizadas no dia-a-dia, como a umidade medida em uma determinada cidade diariamente, ou o preço de uma saca de uma determinada *commodity*, ou a cotação do dólar em relação ao real. MORETTIN (2018) define processo estocástico:

#### Definição 2.1 (Processo estocástico)

Seja  $T$  um conjunto arbitrário e totalmente ordenado e seja  $S$  um conjunto de estados possíveis. Um processo estocástico é uma família de variáveis aleatórias  $Z = \{Z(t) : t \in T\}$ .

#### Definição 2.2 (Autocovariância)

Seja  $Z(t)$  um processo estocástico e sejam  $t_1, t_2 \in T$ . A função de autocovariância  $\gamma(t_1, t_2)$  de  $Z$  é dada por:

$$\gamma(t_1, t_2) = E[Z(t_1)Z(t_2)] - E[Z(t_1)]E[Z(t_2)].$$

#### Definição 2.3 (Autocorrelação)

Seja  $Z(t)$  um processo estocástico e sejam  $t_1, t_2 \in T$ . A função de autocorrelação  $\rho(t_1, t_2)$  de  $Z$  é dada por:

$$\rho(t_1, t_2) = \frac{\gamma(t_1, t_2)}{\sqrt{\gamma(t_1, t_1)\gamma(t_2, t_2)}}.$$

Para algumas séries temporais qualquer, as funções de autocorrelação e autocovariância podem se tornar muito complexas, por isso, em geral restringimos nossos estudos as séries estacionárias.

### Definição 2.4 (Processo estacionário)

Um processo  $Z(t)$  é dito estacionário se suas características são invariantes no tempo, ou seja:

$$F_Z(t) = F_Z(t + h),$$

em que  $F_Z(T)$  é a função de distribuição acumulada do processo  $Z(t)$ .

Assumir que um processo estocástico é estacionário é uma suposição forte que em geral não é atendida, para relaxar um pouco essa suposição, introduzimos a noção de estacionariedade fraca: As séries estacionárias possuem várias propriedades interessantes que facilitam seu estudo. No entanto, não é incomum encontrar processos não estacionários e para contornar estes problemas, existem algumas técnicas capazes de transformar as séries temporais em séries estacionárias.

A primeira técnica é conhecida como diferenciação, quando  $E[Z(t)]$  é da forma  $E[Z(t)] = \mu_0 + \mu t$ , então o processo  $Z^*(t) = Z(t) - Z(t - 1)$  possui esperança  $\mu$  que não depende de  $t$ . A figura 2.1 apresenta o efeito da diferenciação em uma série temporal simulada:

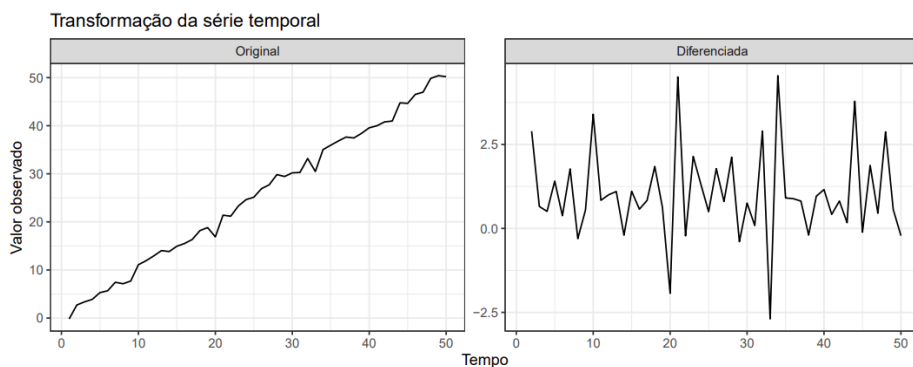


Figura 2.1: À esquerda a série original e à direita a série diferenciada

Note como no primeiro gráfico a série temporal apresenta uma tendência crescente,

enquanto no segundo a série oscila em torno de um valor constante.

Também é comum encontrar séries em que a variância muda ao longo do tempo, nesses casos, utilizar  $Z^*(t) = \log(Z(t))$  pode estabilizar a variância. A figura 2.2 apresenta a transformação logarítmica aplicada na série temporal “Air Passengers”, esta série mede a quantidade de passageiros internacionais de avião por mês entre 1949 e 1960 WOODWARD (2017).

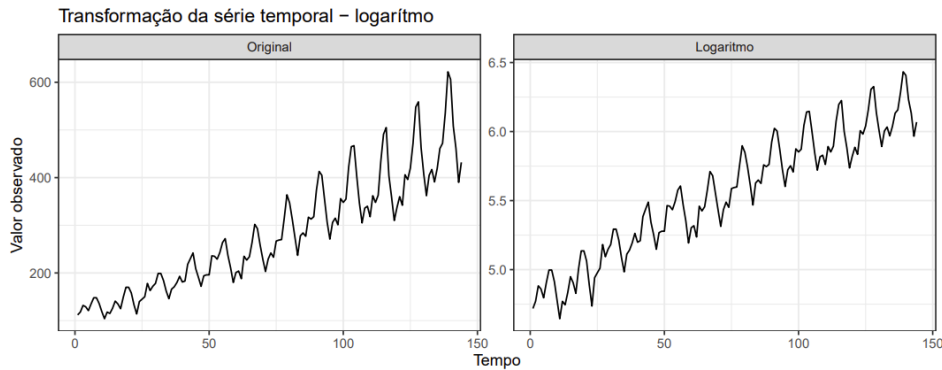


Figura 2.2: A série temporal  $Z(t)$  à esquerda e a série estabilizada  $\log(Z(t))$  à direita.

Uma vez que o logaritmo é uma função estritamente crescente, tomar o logaritmo da série não faz com que a variância seja exatamente constante ao longo do tempo, no entanto, faz com que as diferenças nas variâncias sejam mais sutis, nos permitindo supor que a série é estacionária.

Existem algumas outras técnicas que podem ser utilizadas para obter os mesmos resultados (ex: transformação de Box-Cox). O leitor pode conferir algumas dessas técnicas no terceiro capítulo de HYNDMAN (2018).

## 2.2 Wavelets

Segundo Percival (2000) *wavelets* são ferramentas matemáticas para se analisar séries temporais e imagens. Neste estudo será utilizado para séries temporais, uma sequência ordenada de observação associada a uma variável independente  $t$ . A variável  $t$  pode assumir um conjunto de valores discretos ou contínuos, como os números que pertencem aos  $\mathbb{R}$ .

As técnicas de análise *wavelets* são complementares às técnicas de correlação e análise espectral. Podem ser ferramentas poderosas para se resolver problemas sem soluções satisfatórias antes do surgimento de análise via *wavelet*.

A figura 2.3 apresenta as duas *wavelet* principais:

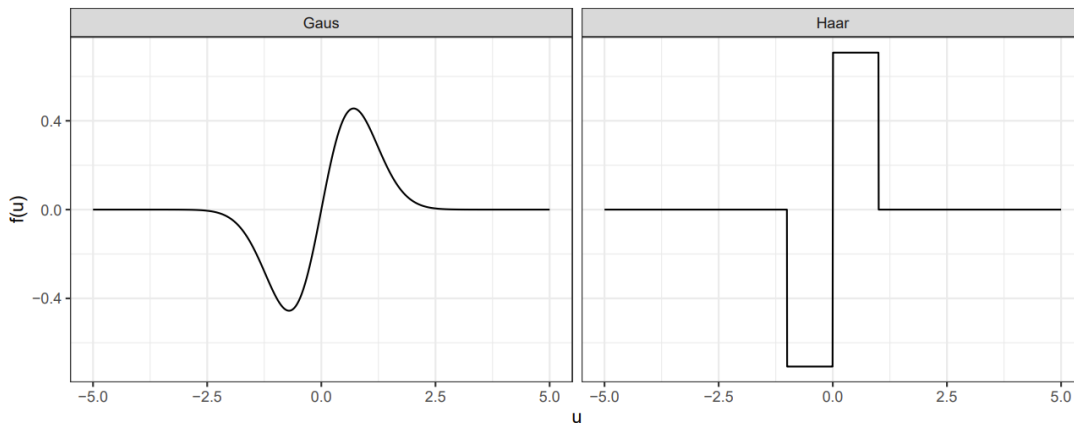


Figura 2.3: À esquerda a *wavelet* de Gauss e à direita a *wavelet* de Haar.

A primeira onda é conhecida como a transformada *wavelet* contínua (TWC), e é utilizada em séries temporais definidas ao longo do eixo real. A segunda é a transformada discreta *wavelet* (TWD), que é aplicada a séries definidas sobre o conjunto de índices  $t = 0, 1, \dots, N - 1$ , onde  $N$  é o número de observações da série temporal.

As funções *wavelets* são utilizadas para gerar bases ortonormais com propriedades interessantes em  $L^2$ , que é o espaço de todas as funções quadrado-integráveis, ou seja:

$$L^2 = \left\{ f : \mathbb{R} \mapsto \mathbb{R} \int_{-\infty}^{\infty} f(x)^2 dx < +\infty \right\}. \quad (2.5)$$

O método de *wavelet* consiste em representar sinais em diferentes escalas de frequência e tempo. Tal sinal pode ser decomposto por uma base formada por *wavelets* transladadas de diferentes escalas e, assim, se obtém detalhes da sua variabilidade ao longo do tempo.

Esta transformação apresenta uma vantagem em relação a outras transformações como a de Fourier, pois apresenta um melhor desempenho para análises em situações com perturbações repentinas, como uma queda repentina nas ações de uma empresa após a pandemia ou a alta inesperada da cotação de criptomoedas. Estes exemplos caracterizam séries temporais não-estacionárias, onde *wavelets* são mais recomendadas porque o método é baseado na extração de informações sobre a variabilidade para diferentes momentos de tempo ou diferentes escalas através dos coeficientes de *wavelets*.

### 2.2.1 Funções *wavelets*

Com uma função *wavelet* que geralmente é representada por  $\psi$  as seguintes propriedades Barbosa e Blitzkow (2008):

1. A integral da função *wavelet* deve ser zero. Isso garante que a função *wavelet* tenha uma do tipo onda. Essa condição é conhecida como condição de admissibilidade. A partir da admissibilidade podemos obter a “Transformada Inversa da função *wavelet*” 2.6.

$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (2.6)$$

2. A função *wavelet* deve ter energia unitária. Isso garante que a função *wavelet* possua suporte compacto, ou com um decaimento rápido de amplitude, garantindo a locação temporal Barbosa e Blitzkow (2008) 2.7.

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1. \quad (2.7)$$

As propriedades mencionadas anteriormente sugerem que a função  $\psi(t)$  apresenta oscilações acima e abaixo do eixo  $t$ , e possui sua energia localizada em uma determinada região, por ser finita. Essa característica de energia concentrada em uma região finita é atribuída a análise de sinais utilizando a Transformada *wavelet*, ainda que utilize funções trigonométricas, como seno e cosseno, que são periódicas e, portanto, possuem energia infinita.

Agora, segundo Percival (2000) a equação 2.7 vale, então para qualquer  $\epsilon$  que satisfaça  $0 < \epsilon < 1$ , deve existir um intervalo  $[-U, U]$  de duração tal que

$$\int_{-U}^U \psi(t)^2 dt < 1 - \epsilon, \quad (2.8)$$

se  $\epsilon$  encaminhar-se para zero, então  $\psi(\cdot)$  só desviaria significativamente de zero fora do intervalo  $[-U, U]$ , pois para valores diferentes de zero a sua resposta fica limitada ao intervalo finito  $[-U, U]$ . O tamanho de  $[-U, U]$  é pequeno em relação ao comprimento infinito de todo o eixo real  $(-\infty, \infty)$ , logo, para valores diferentes de zero em  $\psi(\cdot)$ , podem ser consideradas limitadas a um intervalo de tempo relativamente pequeno. Enquanto a equação 2.7 diz que  $\psi(\cdot)$  tem que fazer algumas excursões para longe de zero, a equação 2.6 nos diz que qualquer excursão que ela faça acima de zero deve ser cancelada por excursões abaixo de zero, então  $\psi(\cdot)$  deve se assemelhar a uma onda. Portanto, as equações 2.6 e 2.7 levam a uma pequena onda ou *wavelet*.

Por exemplo, a *wavelet* mãe de Morlet consiste de uma onda plana modulada por uma

função Gaussiana que é expressa por

$$\psi(t) = \pi^{-\frac{1}{4}} \left( e^{i\xi t} - e^{-\frac{\xi^2}{2}} \right) e^{-\frac{t^2}{2}} \quad (2.9)$$

em que  $\xi$  é uma frequência não-dimensional, plotada na Figura 2.4(a), apresentada mais adiante. Em geral, escolhe-se  $\xi = 5$ , de tal forma que o maior e o menor valor de  $\psi$  seja  $\approx \frac{1}{2}$ . Com isso, a condição de admissibilidade é satisfeita Daubechies (1992). Outra *wavelet* mãe, conhecida como chapéu mexicano, é a derivada segunda da função de densidade de probabilidade Gaussiana, expressa por

$$\psi(t) = \frac{2}{\sqrt{3}\pi^{\frac{1}{4}}} e^{-\frac{t^2}{2}} (1 - t^2) \quad (2.10)$$

Essa função possui suporte infinito. Entretanto, o seu suporte efetivo está no intervalo  $[-5, 5]$ , como pode ser observado na Figura 2.4(b).

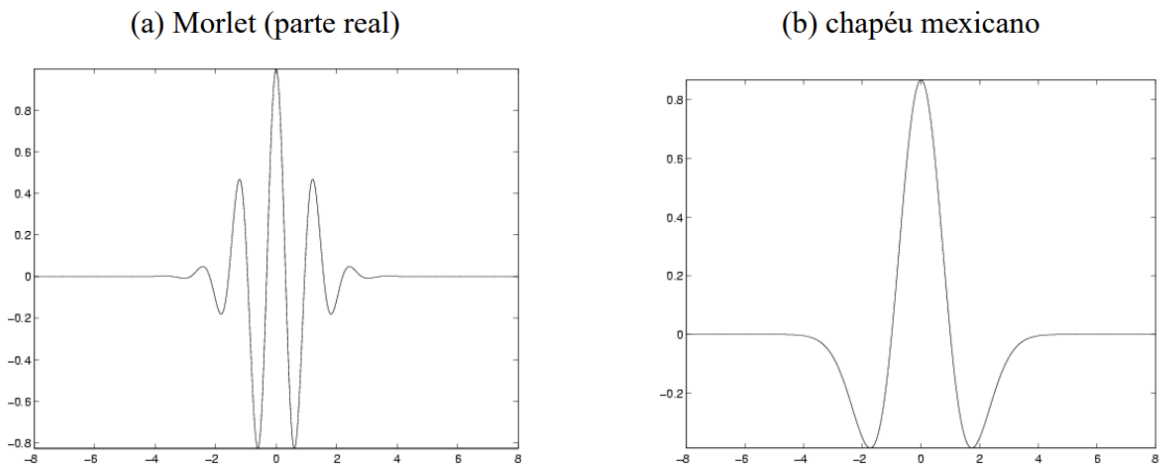


Figura 2.4: Exemplos de funções *wavelet* mãe usadas na TCW.

## 2.2.2 Transformação *Wavelet* Contínua (TWC)

A transformada *wavelet* é uma transformada linear que pode ser utilizada na análise de sinais não estacionários para extrair informações das variações em frequência desses sinais e para detectar suas estruturas temporalmente e/ou espacialmente localizadas.

Pela figura 2.5 temos um esquema de como é aplicada a transformação em um sinal qualquer. À esquerda são apresentadas variações proporcionais dos intervalos de tempo e de frequência e à direita uma representação da dilatação de uma função *wavelet* nesses respectivos intervalos.



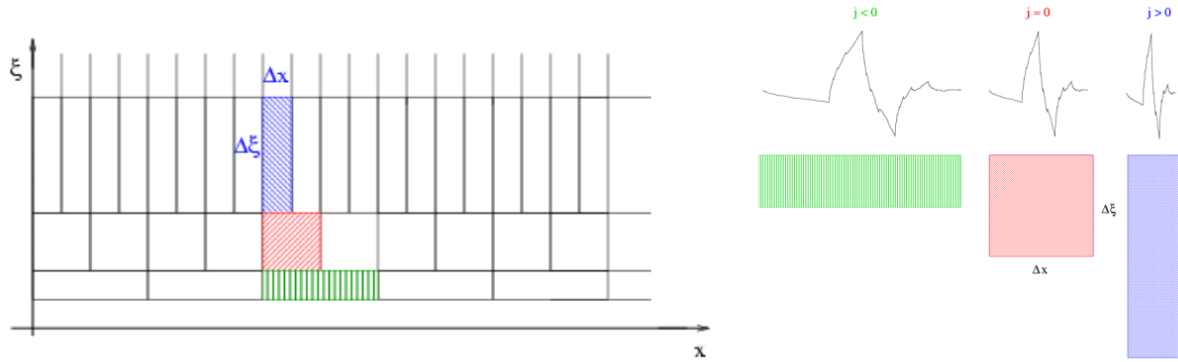


Figura 2.5: Esquema de um plano tempo  $\times$  frequência.

A transformada wavelet de uma série temporal  $f$  é definida pela transformada integral

$$W(a, b) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi * \left( \frac{t-b}{a} \right) dt, \quad (2.11)$$

que é uma função de dois parâmetros reais,  $a$  e  $b$ . O símbolo  $*$  indica o conjugado complexo. Se definirmos  $\psi_{a,b}(t)$  como:

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \psi * \left( \frac{t-b}{a} \right). \quad (2.12)$$

Podemos reescrever a transformada como o produto interno das funções  $f(t)$  e  $\psi_{a,b}(t)$ :

$$W(a, b) = \langle f(t), \psi_{a,b}(t) \rangle = \int_{-\infty}^{\infty} f(t) \psi_{a,b}(t) dt. \quad (2.13)$$

A função  $\psi(t)$  que equivale a  $\psi_{1,0}(t)$  é chamada de "wavelet mãe" (em inglês: mother wavelet) enquanto as outras funções  $\psi_{a,b}(t)$  são chamadas de "wavelets filhas". Estas têm a mesma forma geral que a wavelet mãe.

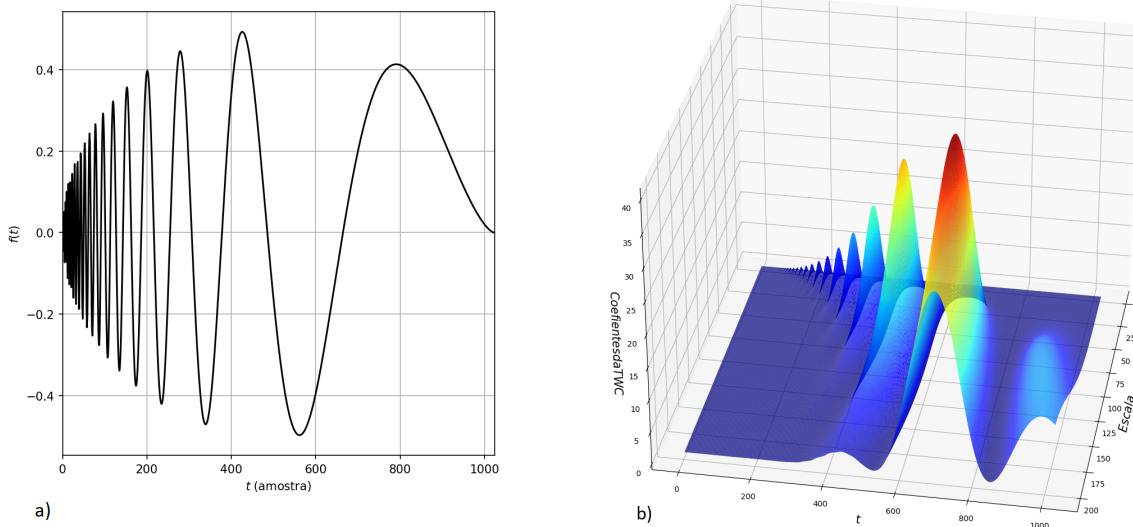
O parâmetro  $b$  indica que a função  $\psi(t)$  foi transladada no eixo  $t$  de uma distância equivalente a  $b$ , sendo então um parâmetro de translação. Já o parâmetro  $a$  causa uma mudança de escala, aumentando (se  $a > 1$ ) ou diminuindo (se  $a < 1$ ) a wavelet formada pela função. Por isto o parâmetro  $a$  é conhecido como parâmetro de escala (em inglês: scaling parameter). As wavelets filhas constituem, por conseguinte, uma família de curvas com forma idêntica à da wavelet mãe, deslocadas no tempo e escaladas em amplitude. No domínio do tempo, a transformada de wavelet é uma medida da correlação entre o sinal  $f(t)$  e as wavelets filhas.

O termo  $\frac{1}{\sqrt{|a|}}$  é um fator de normalização que garante que a energia de  $\psi_{a,b}(t)$  seja

independente de  $a$  e de  $b$ , tal que:

$$\int_{-\infty}^{\infty} |\psi_{a,b}(t)|^2 dt = \int_{-\infty}^{\infty} |\psi(t)|^2 dt. \quad (2.14)$$

Para exemplificar a aparência dessa transformada, na Figura 2.6(a) é apresentado o sinal de Doppler. Na Figura 2.6(b) mostra-se uma representação da transformada desse sinal. Nessa figura observa-se que pequenas escalas correspondem às frequências mais altas, e as frequências decaem à medida que as escalas aumentam. No início desse sinal ocorre uma oscilação de baixa frequência, que está representada nas pequenas escalas e nos pontos de translação de 0 a 300. A seguir há uma oscilação de 0,1 a 0,15, que é a maior frequência do sinal, além de apresentar o pico de maior amplitude, e assim por diante.



*Fonte:* Elaboração Própria.

Figura 2.6: Representação da transformada *wavelet* de um sinal; (a) Série da frequência da aceleração; (b) visualização 3D.

### 2.2.3 Transformação *Wavelet* Discreta (TWD)

De acordo com [Aguiar-Contraria \(2011\)](#) a TWD é derivada da transformada contínua fazendo-se  $a$  e  $b$  variáveis discretas, em lugar de contínuas, através de uma amostragem. O usual é tomar  $a = N^j$ ;  $b = kMa$ ;  $j, k \in \mathbb{N}$

A *wavelet* mãe resultante é uma grade (matriz) de valores  $h[j, k]$  relacionada à *wavelet*

contínua  $\psi(t)$  pela expressão

$$h[j, k] = N^{-\frac{j}{2}} \psi \left( \frac{t}{N^j} - kM \right). \quad (2.15)$$

Quanto mais próximo de 1 o valor de  $N$ , mais a versão discreta se aproxima da versão contínua.  $N$  normalmente é feito igual a 2 e  $M$  igual a 1, mas para agilizar os cálculos em computadores e obter-se um número de coeficientes  $h_{i,j}$  não muito grande, e ainda conservando toda a informação original de  $f(t)$ .

De forma similar, a transformada discreta é uma matriz de coeficientes  $W_{j,k}$  dados por

$$W[j, k] = \sum_{k=1}^m \sum_{j=1}^n f[k] h^* [j, k] = \langle f, h \rangle, \quad (2.16)$$

onde o asterisco denota o conjugado complexo e  $f[k]$  é a sequência de amostras obtida de  $f(t)$ . A distância entre os pontos da grade  $W$  não é uniforme, refletindo a propriedade de auto-escalamento da transformada de *wavelet* contínua. Os limites  $n$  e  $m$  são determinados indiretamente pela escolha de  $N$  e  $M$ .

## 2.3 Como escolher uma função *wavelet*

Não existe uma receita única para escolher uma função *wavelet* que seja mais adequada para analisar um determinado sinal. Por outro lado, algumas recomendações podem ser úteis.

A forma da função *wavelet* escolhida deve refletir as características da série temporal. Por exemplo, para as séries temporais com variações bruscas ou degraus, a *wavelet* de Haar pode ser a mais adequada; por outro lado, para analisar séries temporais com variações mais suaves, as *wavelet* de Morlet e chapéu mexicano são as mais utilizadas.

Para estudar mudanças de amplitude e fase, uma *wavelet* complexa pode ser a mais adequada, como a de Morlet. Isso ajuda a capturar o comportamento oscilatório dos dados.



## Capítulo 3

# Correlação entre séries temporais via *wavelet*

O conceito de correlação é o mesmo usado em dados que são de séries temporais: identificar e quantificar a relação entre duas variáveis quantitativas. Devido à natureza contínua e cronologicamente ordenada dos dados das séries temporais, é provável que haja algum grau de correlação entre as observações das séries.

Ao analisar a correlação entre uma série e seus *lags* se pode encontrar informações preditivas, que podem ser utilizadas para prever eventos da série. Os métodos mais populares para medir o nível de correlação entre uma série e suas defasagens são a função de autocorrelação (FAC) e a função de autocorrelação parcial (FACP).

Medir e analisar a correlação entre duas variáveis, no contexto da análise de séries temporais pode ser entendido por: analisar a correlação entre duas séries para identificar fatores exógenos ou preditores, que possam explicar a variação da série ao longo do tempo. Nesse caso, a medição da correlação é normalmente feita com a função de correlação cruzada.

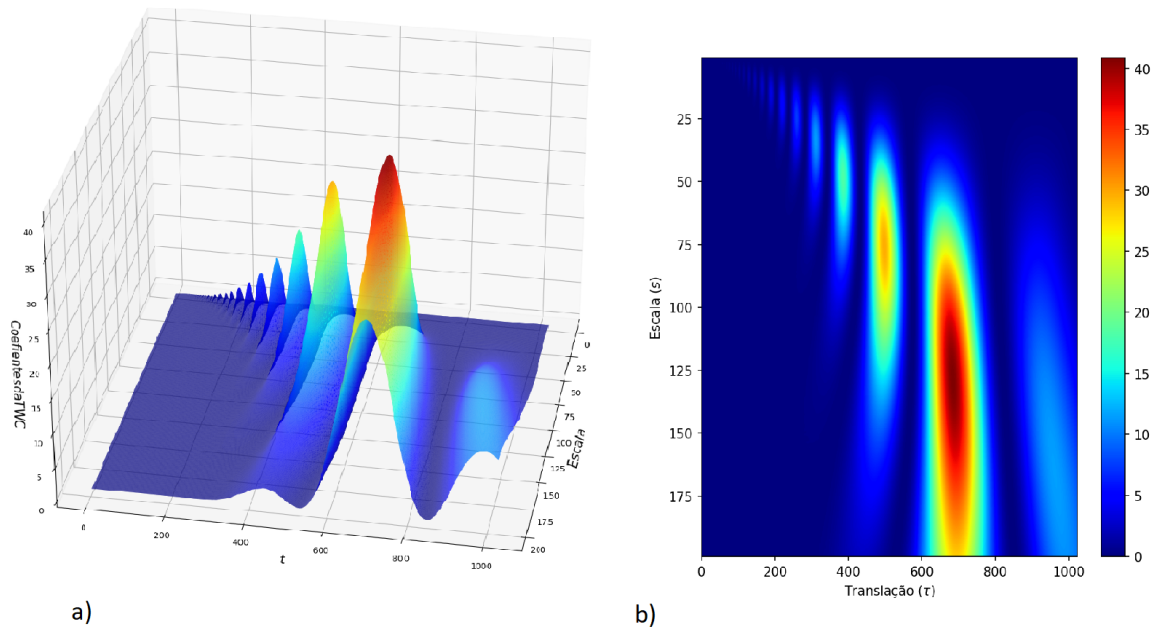
Neste estudo o interesse é encontrar outra maneira para se medir a correlação entre séries temporais.

De acordo com [K. Gurley e A. Kareem \(2002\)](#), os coeficientes *wavelet* calculados são adequados para analisar sinais não estacionários, com seus valores quadrados plotados em uma grade de escala de tempo.

Para isto existe a visualização de dados chamada de escalograma ou mapa quadrado médio, que mostra o conteúdo da frequência do sinal em cada intervalo de tempo para identificar a ocorrência de transientes, enquanto acompanha os fenômenos evolutivos no

tempo e na frequência. Segundo [K. Gurley e A. Kareem \(2002\)](#) o conceito de escalograma foi avançado para identificar a correlação entre sinais, nos quais os coeficientes quadrados são substituídos pelo produto dos coeficientes *wavelet* de dois processos diferentes. Este coscalograma produz uma visão dos eventos coincidentes entre os processos, revelando bolsões de correlação ao longo da frequência que variam no tempo.

Exemplo disto temos na [Figura 3.1\(b\)](#), onde se pode observar a representação tridimensional da TWC da [Figura 3.1\(a\)](#) no escalograma ao lado.



Fonte: Elaboração Própria.  
 Figura 3.1: Representação da transformada *wavelet* de um sinal; (a) visualização 3D; (b) visualização pelo escalograma

Deve-se notar que outros estudos aplicaram análise espectral de ordem superior para quantificar a relação não linear entre a velocidade do vento e a pressão [Gurley e Kareem \(1997\)](#). No entanto, esses métodos espectrais de ordem superior baseados em Fourier não são capazes de capturar a relação intermitente transiente que está sendo procurada aqui. Uma abordagem mais precisa e confiável para identificar quantitativamente correlações intermitentes de primeira ordem e de ordem superior é o objetivo deste estudo.

### 3.1 Mapa de coerência de *wavelet*

A Coerência *Wavelets* se parece com coeficiente de correlação aplicado em diferentes momentos do tempo entre dois sinais.

Segundo [GURGEL \(2013\)](#), os coeficientes da Coerência *wavelets* são visualizados em um mapa de cores (escalograma) para dois sinais quaisquer. O eixo horizontal representa o domínio temporal e o vertical as escalas utilizadas para realizar o cálculo da Coerência *wavelets*. Por [GRINSTED \(2003\)](#), pode-se entender que as regiões com baixa coerência coincidem com baixo espectro de energia das *wavelets*. Sendo que os espectros de baixa energia são esperados devido a não captura da localização real dos centros de ação, entre duas correlações. Esta técnica de Coerência para correlacionar variáveis diferentes em um

mesmo período mostram as bandas de frequência dentro das quais duas séries temporais estão se correlacionando.

A forma tradicional da função de coerência é definida como a razão do espectro cruzado de *wavelet*, para o produto dos auto espectros de *wavelets* de dois sinais  $x(t)$  e  $y(t)$ . O mapa de coerência *wavelet* é definido como

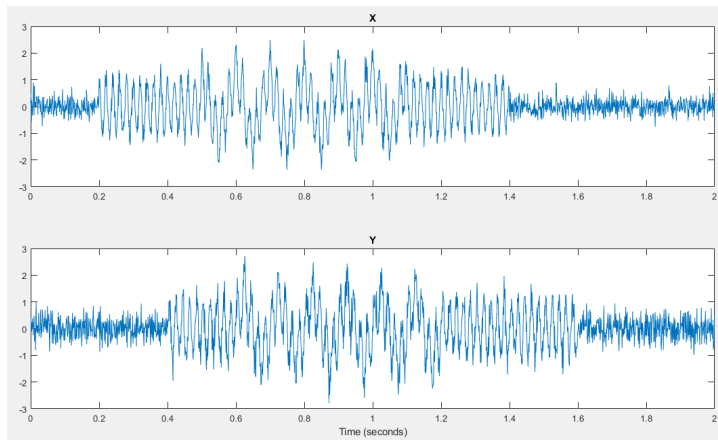
$$[c^W(a, t)]^2 = \frac{|S_{xy}^W(a, t)|^2}{S_{xx}^W(a, t)S_{yy}^W(a, t)}, \quad (3.1)$$

onde os espectros de potência localizados, discutidos acima, são dados por

$$S_{ij}^W(a, t) = \int_T W_i^*(a, t)W_j(a, t)dt. \quad (3.2)$$

A janela de integração de tempo localizada em 3.2,  $T = [t - \Delta T, t + \Delta T]$ , é selecionada com base na resolução de tempo desejada no mapa de coerência resultante, é essencialmente executa a mesma operação de média de conjunto, embora localizada no tempo, como a análise de Fourier tradicional para obter um espectro automático ou espectro cruzado de dois sinais. O mapa é limitado entre 0 e 1 e fornece uma visão da correlação localizada em relação ao tempo e à frequência de acordo com K. Gurley e A. Kareem (2002).

Na figura 3.2 temos um exemplo de dois sinais para obter os dados de coerência *wavelet*, especificando um intervalo de amostragem de 0,001 segundos. Ambos os sinais consistem em duas ondas senoidais (10 Hz e 50 Hz) em ruído branco. As ondas senoidais têm suportes de tempo diferentes.



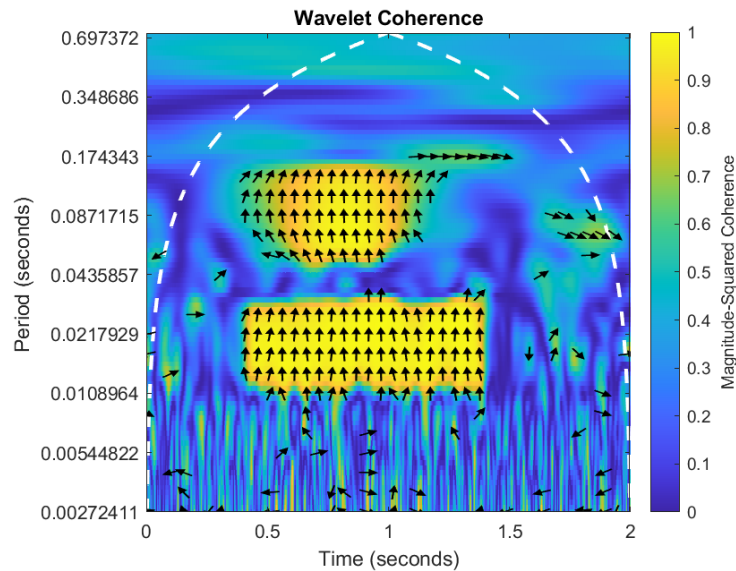
*Fonte:* Elaboração Própria.

Figura 3.2: Gráfico de duas ondas senoidais (10 Hz e 50 Hz) em ruído branco

Devido à relação inversa entre frequência e período, um gráfico que usa o intervalo de



amostragem é o inverso de um gráfico que usa a frequência de amostragem. Para áreas onde a coerência excede 0,5, os gráficos que usam a frequência de amostragem exibem setas para mostrar o atraso de fase de  $y$  em relação a  $x$ . As setas são espaçadas no tempo e na escala. A direção das setas correspondem ao atraso de fase no círculo unitário. Por exemplo, uma seta vertical indica um atraso de fase  $\frac{\pi}{2}$  ou quarto de ciclo. O atraso correspondente no tempo depende da duração do ciclo.



*Fonte:* Elaboração Própria.

Figura 3.3: Gráfico do mapa de coerência entre dois sinais

No escalograma o eixo  $y$  corresponde aos valores dos períodos que decompõem a série pela frequência do seu sinal, ou seja, o quão mais alto é o período mais baixa é a frequência da série e mais alta é a amplitude. Assim o gráfico pode mostrar em se as séries são correlacionadas em tendência e variabilidade. Pelo gráfico 3.3 percebemos que entre os 0.5 e 1.5 segundos as série é altamente correlacionada para frequências médias.



# Capítulo 4

## Aplicação

O objetivo do estudo é realizar uma análise descritiva e multivariada utilizando técnicas de transformação *wavelet* contínuas e discretas para se analisar a correlação entre determinados sinais. E esses sinais que serão utilizados no estudo são o BOVESPA, *Dow Jones Industrial Average*, Bitcoin e S&P500.

### 4.1 Banco de dados

O banco de dados contém os dados diários de Janeiro de 2012 até Julho de 2023 dos preços de fechamento dos ativos financeiros IBOVESPA, DJIA, Bitcoin e S&P500 que foram obtidas através do *yfinance*, que é uma API dedicada, que faz uso da base de dados do *Yahoo Finance*.

O Índice Bovespa (BOVESPA), é o mais importante indicador do desempenho médio das cotações das ações negociadas na B3. É formado pelas ações com maior volume negociado nos últimos meses. Trata-se da formação de uma carteira de investimentos teórica que, ao final de 2021, era composta de 91 ações de 84 empresas, retratando a movimentação dos principais papéis negociados na B3, representando não só o comportamento médio dos preços mas também o perfil das negociações observadas nos pregões.

O *Dow Jones Industrial Average* (DJIA) é um dos principais indicadores dos movimentos do mercado norte-americano. O cálculo deste índice é baseado na cotação das ações de 30 das maiores e mais importantes empresas dos Estados Unidos.

O Bitcoin surgiu em 2008 como uma resposta à crise financeira, com a ideia de substituir o dinheiro físico que usamos e, principalmente, tirar a necessidade de intermediação dos bancos nas operações financeiras. A criptomoeda apareceu pela primeira vez em um

artigo publicado por [Nakamoto \(2008\)](#). O texto descrevia como funcionava esta moeda digital e criava o sistema que posteriormente passou a ser chamado de “blockchain”, que é como um livro-razão que registra todas as operações.

Além dessas ações, temos o S&P500, que é uma abreviação de *Standard & Poor's 500*, ou simplesmente S&P, que se trata de um índice composto por quinhentos ativos (ações) cotados nas bolsas de NYSE ou NASDAQ, qualificados devido ao seu tamanho de mercado, sua liquidez e sua representação de grupo industrial.

Por fim, existe uma justificativa para a escolha de tais séries financeiras. A série do *Dow Jones* pode ter uma forte correlação positiva com o S&P500, já que eles têm empresas semelhantes em suas composições. O *Dow Jones* com o Bitcoin pode ser fracamente correlacionados, por conta da sua independência do mercado financeiro, pois os Bitcoins é uma criptomoeda não lastreada, o que significa dizer que o seu valor não está atrelado a uma reserva financeira ou a um bem físico. O seu valor decorre, principalmente, da lei da oferta e da procura. Já o *Dow Jones* e o BOVESPA podem ter uma correlação positiva, porque eventos que afetam o *Dow Jones* podem influenciar também o mercado brasileiro de ações. Por exemplo: a aceleração da inflação nos Estados Unidos e o aumento da taxa básica de juros por lá, com perspectiva de diminuição da atividade econômica. Este cenário tende a deprimir as bolsas nos EUA e se reflete também no Brasil.

## 4.2 Procedimento metodológicos

Para conhecermos melhor nosso banco de dados, faremos uma breve análise descritiva do mesmo a fim de observar o comportamento das séries temporais em estudo. Logo após, vamos utilizar técnicas de transformação *wavelet*, como decomposição ARM, correlação multiescala e a correlação cruzada via *wavelet*, nas séries financeiras escolhidas para verificarmos o quão correlacionadas elas são entre si. Lembrando que temos informações a respeito de mais de 11 anos das cotações das ações citadas na seção anterior nos anos de Jan/2012 a Jul/2023.

## 4.2.1 Análise descritiva

### Dados Faltantes

Vamos iniciar a nossa análise verificando a quantidade de dados faltantes no nosso banco de dados, logo, vamos utilizar o gráfico de matriz. Ele é uma ótima ferramenta para se trabalhar com dados de séries temporais, pois fornece um preenchimento de cor para cada coluna. Quando os dados existem, o gráfico é sombreado em cinza (ou na cor de sua escolha) e, quando não há dados, o gráfico exibi brancos.

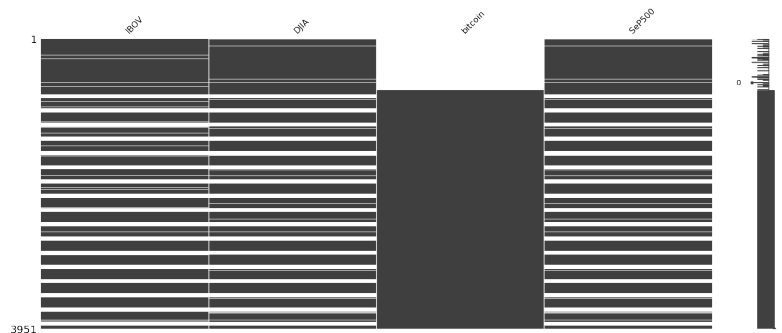


Figura 4.1: Gráfico da matriz de dados faltantes

Conforme visto no gráfico 4.1, as colunas IBOV, DJIA e S&P500 mostram pequenas porções de dados ausentes. E para a coluna do Bitcoin, temos dados ausentes de Janeiro de 2012 até Setembro de 2014. Vale ressaltar que o IBOV, DJIA, Bitcoin e S&P500 apresentam 2869, 2916, 3244 e 2916 observações respectivamente. Isso acontece porque temos datas de feriados diferentes dos EUA e também a API pode ter tido algum problema em relação a coleta desses dados.

Para solucionar isso vamos utilizar uma técnica de substituição de dados faltantes chamada de interpolação linear. Este método determina os valores das funções de quaisquer pontos intermediários quando os valores de dois pontos adjacentes são conhecidos. A interpolação linear é basicamente a estimativa de um valor desconhecido que cai dentro de dois valores conhecidos.

No python, o pacote *SciPy* possui vários métodos de interpolação, como o `interpolate`. Dentro desta função, vamos utilizar o método `backfill` que preenche o local do dado faltante baseado no próximo valor numa direção de trás para frente, ou seja, das datas mais recentes para as mais antigas.

Após a interpolação nas séries financeiras escolhidas, o banco de dados contém 3951 observações para cada um dos sinais.

## Análise univariada

Iniciando as análises, teremos como foco a série financeira da BOVESPA.

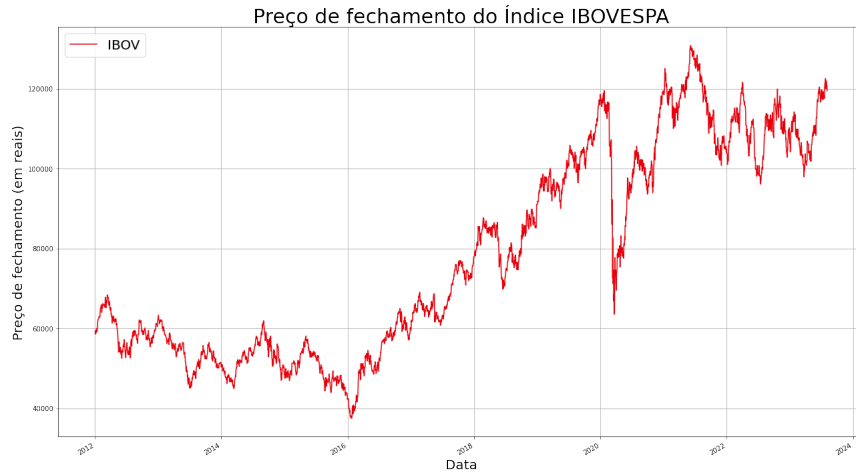


Figura 4.2: Gráfico do preço de fechamento do índice IBOVESPA de 2012 até 2023

No gráfico 4.2 notamos uma aparente tendência decrescente na série entre os anos de 2012 e 2016, por conta de uma recessão econômica depois do boom de 2010. Mas só a partir de 2016 que a BOVESPA fecha o ano em alta, logo depois fica visível a tendência crescente dos preços de fechamento até 2020. E em Janeiro de 2020 até Abril de 2020 verificamos uma queda significativa nos preços por conta da crise causada pela pandemia da COVID-19. A partir de Janeiro de 2021 a série não apresenta tendência, mas uma alta variabilidade.

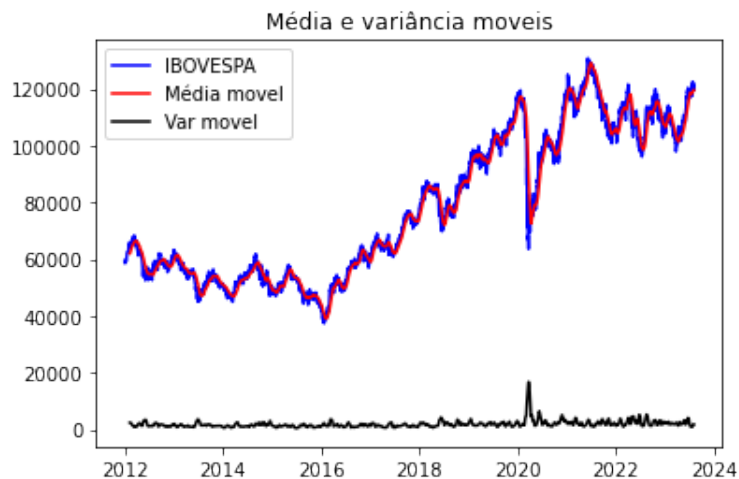


Figura 4.3: Gráfico da média e variância móvel do preço de fechamento do índice IBOVESPA de 2012 até 2023

Também é possível notar que a série não é estacionária, porque na figura 4.3 percebemos que a série não se desenvolve em torno de uma média ou variância constantes. E

pelo teste de Dickey Fuller, o valor-p = 0.7394 >  $\alpha = 0.05$ , então não há evidências de se rejeitar a hipótese nula. Isso significa que a série temporal é não estacionária. Em outras palavras, possui alguma estrutura dependente do tempo e não possui variação constante ao longo do tempo.



Figura 4.4: Gráfico do preço de fechamento do **Dow Jones** de 2012 até 2023

A seguir, vamos analisar a série financeira do *Dow Jones Industrial Average*. No gráfico 4.4 notamos uma aparente tendência crescente ao longo da série. Mas em Dezembro de 2019 houve uma queda significativa dos preços de fechamento, por conta do impacto do Brexit no Reino Unido que preocupou os investidores como também a desaceleração da economia chinesa. E em Janeiro de 2020 até Abril de 2020 verificamos uma queda significativa nos preços por conta da crise causada pela pandemia da COVID-19. Logo após, é visível que existe uma alta significativa nos preços de fechamento até meados de 2022. A partir de Janeiro de 2021 a série não apresenta tendência, mas uma alta variabilidade.

Também é possível notar que a série não é estacionária, porque na figura 4.5 percebemos que a série não se desenvolve em torno de uma média ou variância constantes. E pelo teste de Dickey Fuller, o valor-p = 0.7877 >  $\alpha = 0.05$ , então não há evidências de se rejeitar a hipótese nula. Isso significa que a série temporal é não estacionária. Em outras palavras, possui alguma estrutura dependente do tempo e não possui variação constante ao longo do tempo.

Passemos então à análise da série financeira do Bitcoin. No gráfico 4.6 notamos uma aparente tendência crescente ao longo da série, mas a variabilidade neste caso é muito alta. Mesmo que não tenhamos dados de Janeiro de 2012 até Setembro de 2014 o estado do valor das ações seguiu em baixa e com pouca volatilidade, como na série em estudo.

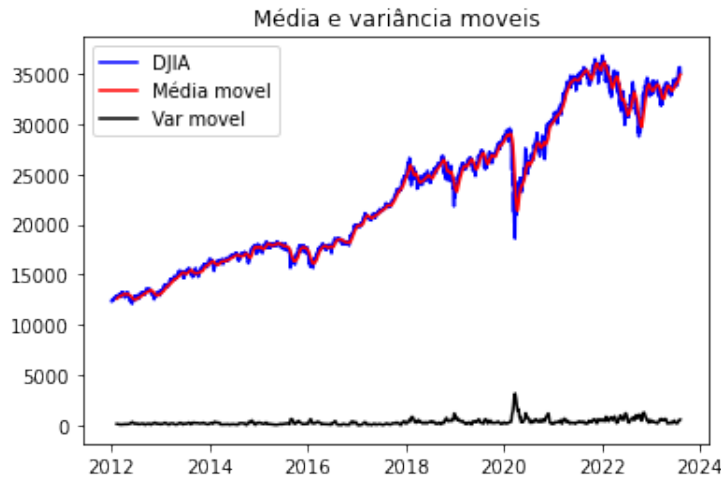


Figura 4.5: Gráfico da média e variância móvel do preço de fechamento do *Dow Jones* de 2012 até 2023

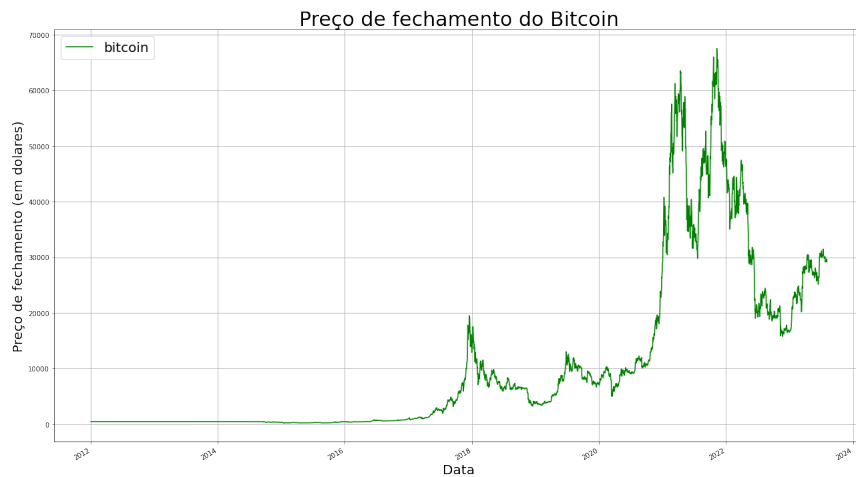


Figura 4.6: Gráfico do preço de fechamento do Bitcoin de 2012 até 2023

O ponto mais significativo desta série são as duas maiores altas nos preços das ações nos meses de Abril e Outubro de 2021, além das maiores quedas que foram nos meses de Maio e Junho de 2022.

Também é possível notar que a série não é estacionária, porque na figura 4.7 percebemos que a série não se desenvolve em torno de uma média ou variância constantes. E pelo teste de Dickey Fuller, o valor- $p = 0.459021 > \alpha = 0.05$ , então não há evidências de se rejeitar a hipótese nula. Isso significa que a série temporal é não estacionária. Em outras palavras, possui alguma estrutura dependente do tempo e não possui variação constante ao longo do tempo.

Agora, seguiremos a análise da série financeira do S&P500. No gráfico 4.8 verificamos uma semelhança muito grande com a série do *Dow Jones*, como a aparente tendência crescente ao longo da série. Ou que em Dezembro de 2019 houve uma queda significativa



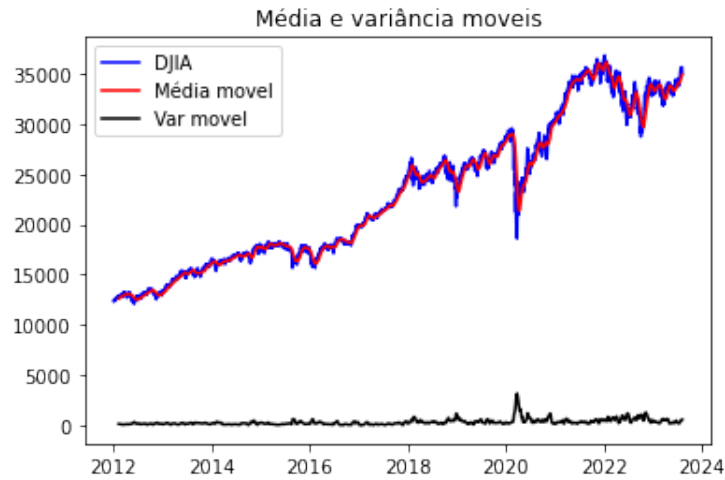


Figura 4.7: Gráfico da média e variância móvel do preço de fechamento do Bitcoin de 2012 até 2023

dos preços de fechamento, por conta do impacto do Brexit no Reino Unido que preocupa os investidores como também a desaceleração da economia chinesa. E em Janeiro de 2020 até Abril de 2020 verificamos uma queda significativa nos preços por conta da crise causada pela pandemia da COVID-19. Logo após, é visível que existe uma alta significativa nos preços de fechamento até meados de 2022. A partir de Janeiro de 2021 a série não apresenta tendência, mas uma alta variabilidade.



Figura 4.8: Gráfico do preço de fechamento do **Dow Jones** de 2012 até 2023

Também é possível notar que a série não é estacionária, porque na figura 4.9 percebemos que a série não se desenvolve em torno de uma média ou variância constantes. E pelo teste de Dickey Fuller, o valor-p = 0.8709 >  $\alpha = 0.05$ , então não há evidências de se rejeitar a hipótese nula. Isso significa que a série temporal é não estacionária. Em outras palavras, possui alguma estrutura dependente do tempo e não possui variação constante

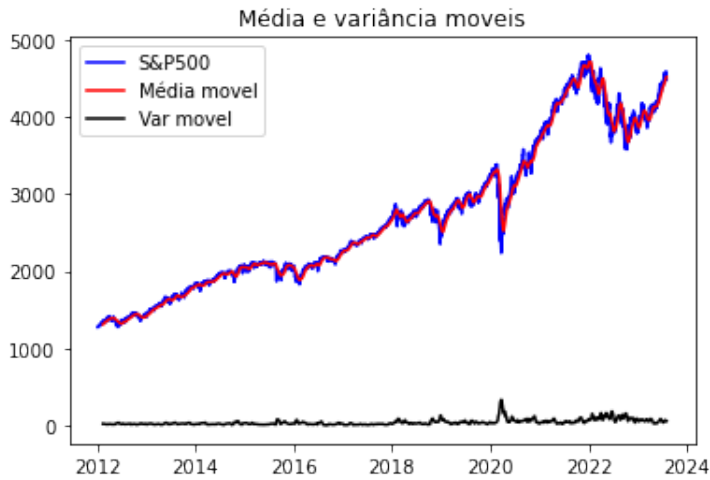


Figura 4.9: Gráfico da média e variância móvel do preço de fechamento do *Dow Jones* de 2012 até 2023

ao longo do tempo.

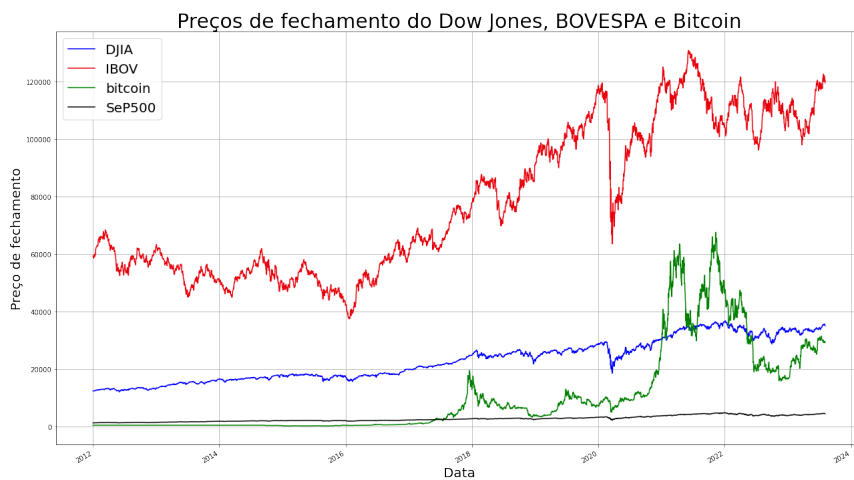


Figura 4.10: Gráfico dos preços de fechamento do *Dow Jones* e IBOVESPA de 2012 até 2023

Finalizando a análise univariada, observamos o conjunto das quatro séries financeiras no gráfico 4.10. Podemos perceber que a maioria delas tem uma tendência de crescimento de 2016 até 2023. Também todas elas sofreram alguma desvalorização no mesmo período da crise causada pelo COVID-19.

Entretanto, algumas séries como o do *Dow Jones* e do BOVESPA apresentam momentos com correlação positiva e em outros momentos correlação negativa, por exemplo Janeiro de 2020 e Outubro de 2022 respectivamente.

Buscando poder descrever melhor a influencia desse fenômeno, vamos utilizar a transformação *wavelet* para compreender melhor a correlação entre séries temporais financeiras.

## Escalograma

Vamos iniciar agora a visualização da transformação *Wavelet* através do escalograma. Em seguida temos essas representações gráficas para cada série financeira estudada.

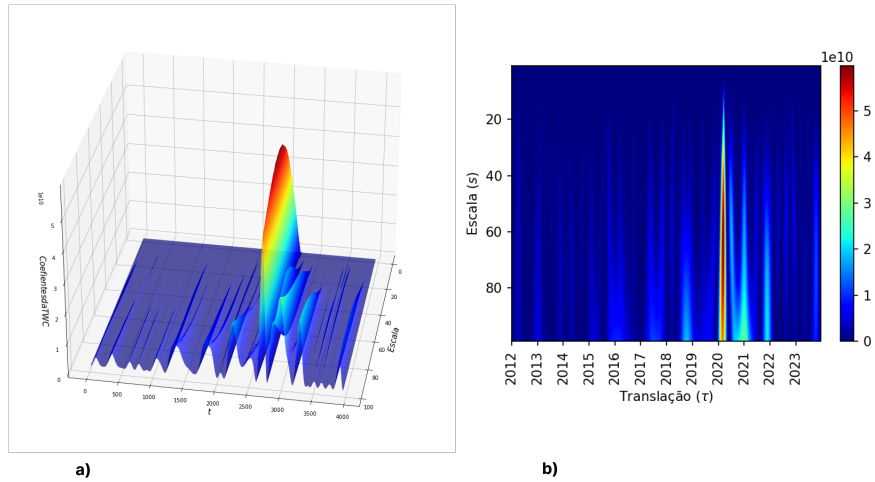


Figura 4.11: Representação da transformada *wavelet* dos preços de fechamento do BOVESPA de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma

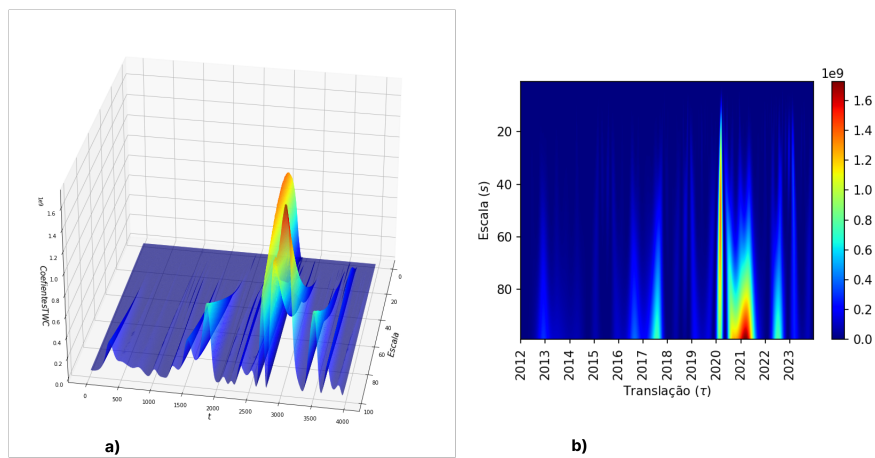


Figura 4.12: Representação da transformada *wavelet* dos preços de fechamento do *Dow Jones* de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma

Perceba que nas figuras 4.11, 4.12, 4.13 e, 4.14 temos picos de frequência no período de crise da pandemia de COVID-19, que de fato causa as maiores deformidades nos gráficos em três dimensões.

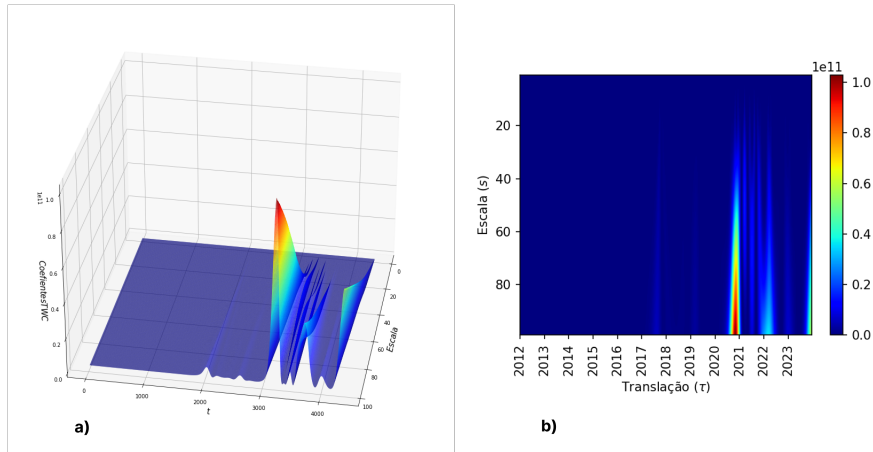


Figura 4.13: Representação da transformada *wavelet* dos preços de fechamento do Bitcoin de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma

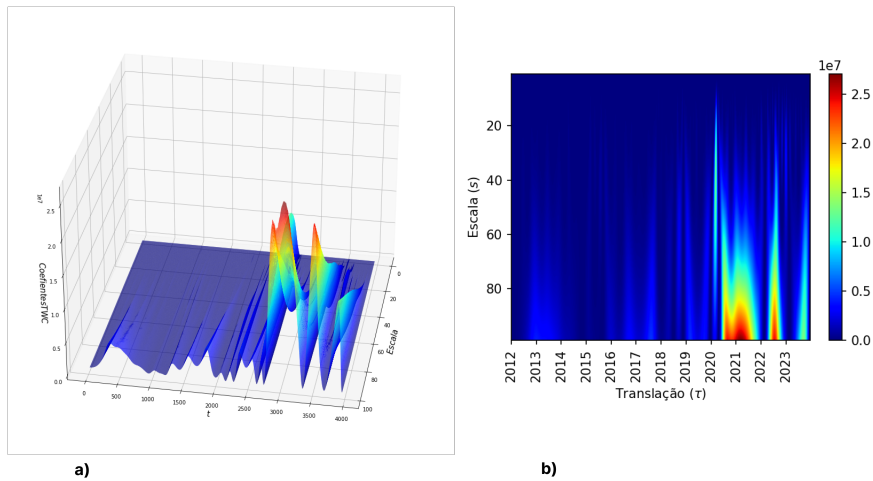


Figura 4.14: Representação da transformada *wavelet* dos preços de fechamento do S&P500 de 2012 até 2023; (a) visualização 3D; (b) visualização pelo escalograma

## Decomposição de séries temporais

Agora vamos realizar uma análise multirresolução de sinal (ARM) da série da Bovespa. Para analisar a transformação *wavelet* aplicada a série. Através desta decomposição vamos ter a capacidade de verificar semelhanças entre as séries através de um filtro, que é a transformação *wavelet*.

Em ambas as decomposições ARM, as oscilações de alta frequência estão localizadas nas primeiras funções de modo intrínseco  $\tilde{D}_1$ . As oscilações de frequência mais baixa estão localizadas em grande parte no  $\tilde{D}_3$ , mas podemos ver algum efeito também no  $\tilde{D}_4$ . Os componentes de tendência estão no  $\tilde{S}_8$ .

Como visto na análise descritiva, temos semelhanças no período de Janeiro de 2020 até Dezembro de 2021. Isto fica mais evidente quando comparamos a tendência das séries

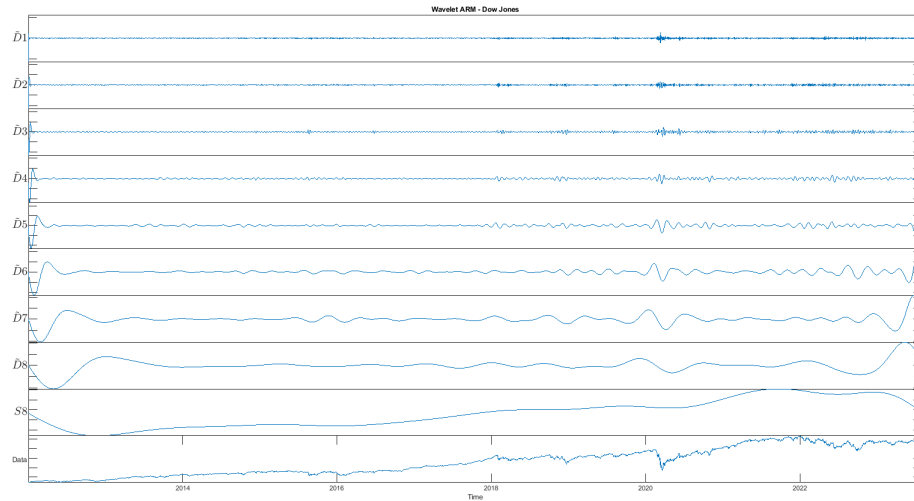


Figura 4.15: Gráfico da **Wavelet ARM** do preço de fechamento do *Dow Jones* de 2012 até 2023

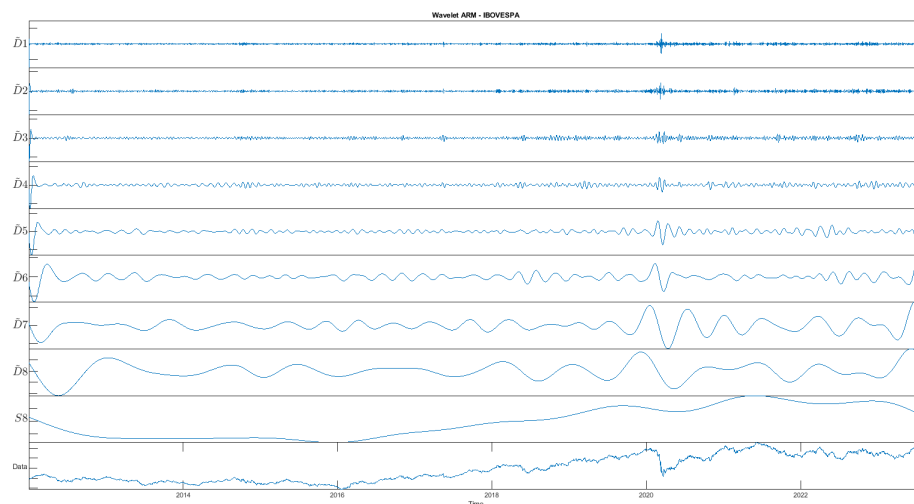


Figura 4.16: Gráfico da **Wavelet ARM** do preço de fechamento do IBOVESPA de 2012 até 2023

nas decomposições de  $\tilde{S}_g$ .

## 4.2.2 Correlação

A correlação multiescala disponível mostra uma correlação negativa significativa para todos os níveis da decomposição. Mesmo essa correlação é apenas marginalmente significativa ao ajustar para comparações múltiplas.

A análise de correlação multiescala revela que as correlações são todas positivas nos dados agregados. A medida em que se aumenta o nível da decomposição, a correlação aumenta. A maior correlação ocorre entre decomposições de tendência.

Com dados financeiros, muitas vezes há uma relação de avanço ou atraso entre as

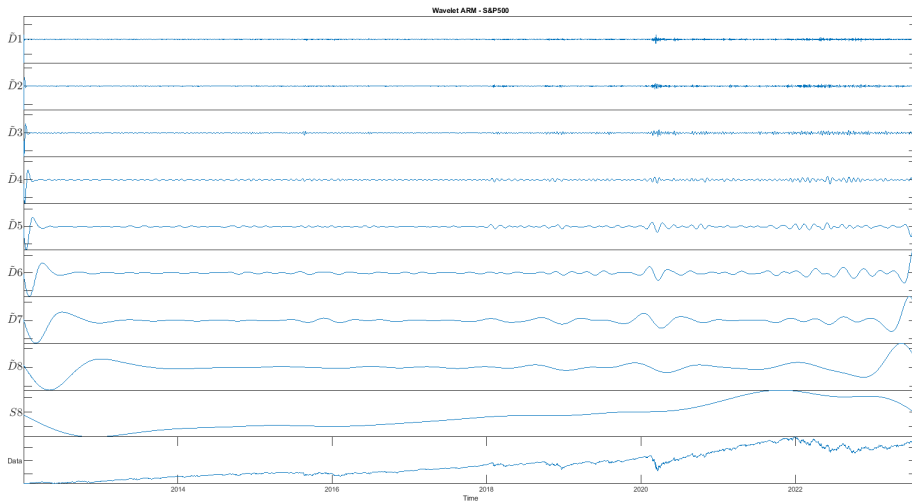


Figura 4.17: Gráfico da **Wavelet** ARM do preço de fechamento do S&P500 de 2012 até 2023

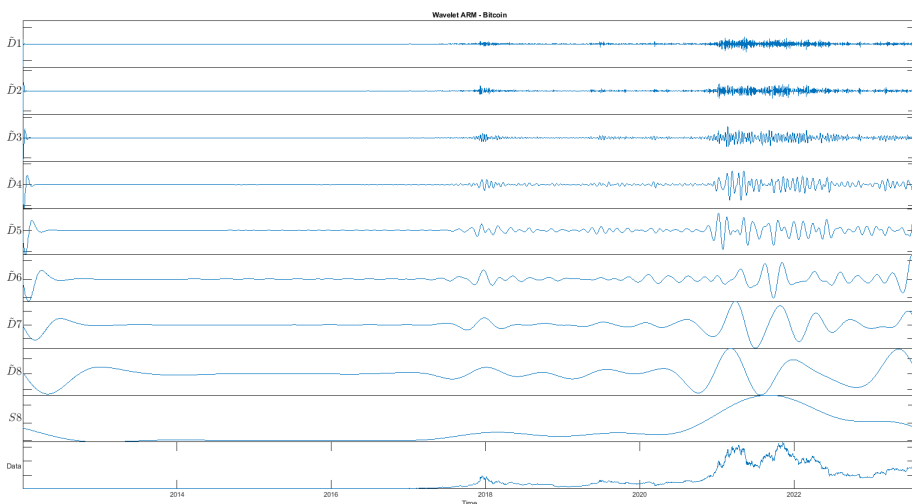


Figura 4.18: Gráfico da **Wavelet** ARM do preço de fechamento do S&P500 de 2012 até 2023

variáveis. Nesses casos, é útil examinar a sequência de correlação cruzada para determinar se o atraso de uma variável em relação a outra maximiza sua correlação cruzada. Percebemos isso no gráfico 4.10 quando mencionamos que algumas séries como o do *Dow Jones* e do BOVESPA apresentam momentos com correlação positiva e em outros momentos correlação negativa.

Na tabela 4.1 temos os resultados da análise da correlação multiescala entre as séries financeiras dos preços de fechamento do IBOVESPA e *Dow Jones*. A representação gráfica dessa correlação está no gráfico 4.19 e nele podemos perceber que em níveis maiores, há uma forte correlação positiva entre os preços de fechamento. Porém os intervalos de confiança para essas correlações vão ficando cada vez maiores à medida que o nível aumenta.

Tabela 4.1: Tabela da correlação multiescala dos preços de fechamento do IBOVESPA e *Dow Jones* de 2012 até 2023

Descomposições	NJ	Superior	Rho	Inferior	Valor-P
D1	3948	0,49204	0,52476	0,55599	0,00000
D2	3942	0,48116	0,52772	0,57130	0,00000
D3	3930	0,48988	0,55422	0,61255	0,00000
D4	3906	0,47666	0,56788	0,64700	0,00000
D5	3858	0,44030	0,57269	0,68071	0,00000
D6	3762	0,34344	0,54788	0,70275	0,00000
D7	3570	0,25394	0,56273	0,76742	0,00121
D8	3186	0,19918	0,64557	0,87009	0,00934
S8	3186	0,88313	0,96075	0,98717	0,00000

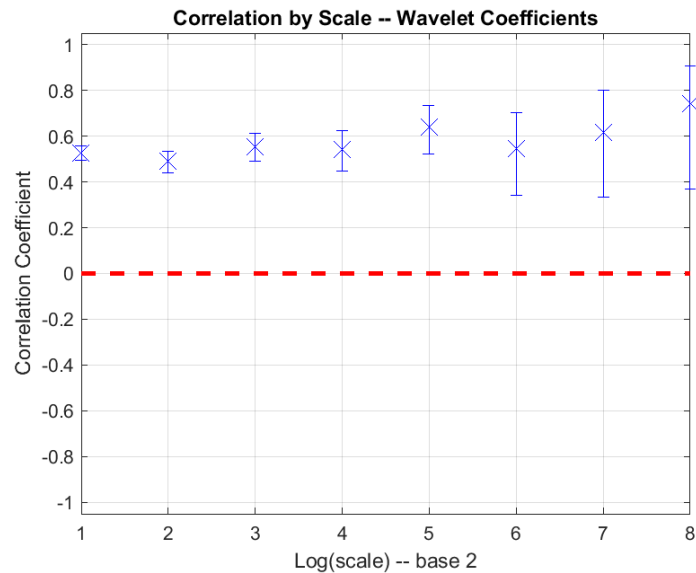


Figura 4.19: Gráfico da correlação multiescala dos preços de fechamento do IBOVESPA e *Dow Jones* de 2012 até 2023

Em seguida na tabela 4.2 temos os dados da correlação multiescala das nossas séries mais correlacionadas, como as séries financeiras dos preços de fechamento do S&P500 e *Dow Jones*.

E a representação gráfica dessa correlação está no gráfico 4.20, onde podemos perceber que em todos os níveis as correlações de preço de fechamento entre essas duas séries, são fortes e positivas. Além disso, os intervalos de confiança para essas correlações são, em sua maioria, pequenos e significativos.

Agora, na tabela 4.3 temos os dados da correlação multiescala das nossas séries menos correlacionadas, como as séries financeiras dos preços de fechamento do Bitcoin e *Dow Jones*.

Tabela 4.2: Tabela da correlação multiescala dos preços de fechamento do S&P500 e *Dow Jones* de 2012 até 2023

Descomposições	NJ	Superior	Rho	Inferior	Valor-P
D1	3948	0,94422	0,94881	0,95303	0,00000
D2	3942	0,93890	0,94588	0,95209	0,00000
D3	3930	0,93447	0,94482	0,95356	0,00000
D4	3906	0,92740	0,94308	0,95545	0,00000
D5	3858	0,91802	0,94196	0,95906	0,00000
D6	3762	0,89244	0,93429	0,96020	0,00000
D7	3570	0,83394	0,91831	0,96073	0,00000
D8	3186	0,77806	0,92260	0,97436	0,00000
S8	3186	0,96255	0,98777	0,99604	0,00000

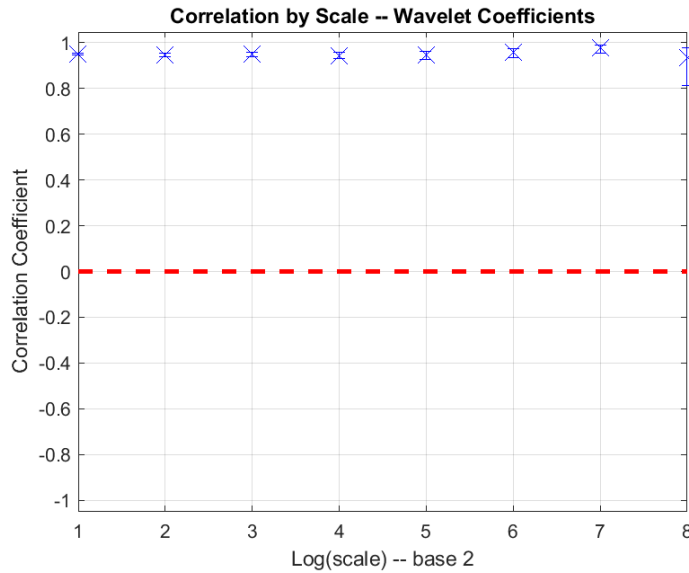


Figura 4.20: Gráfico da correlação multiescala dos preços de fechamento do S&P500 e *Dow Jones* de 2012 até 2023

A representação gráfica dessa correlação está no gráfico 4.21 e nele podemos perceber que em todos os níveis as correlações de preço de fechamento entre essas duas séries, são fracas e positivas. Além disso, os intervalos de confiança para essas correlações vão ficando cada vez maiores à medida que o nível aumenta. Mas só os níveis 7 e 8 contém o zero, mostrando que nessas decomposições as séries não são correlacionadas.



Tabela 4.3: Tabela da correlação multiescala dos preços de fechamento do Bitcoin e *Dow Jones* de 2012 até 2023

Descomposições	NJ	Superior	Rho	Inferior	Valor-P
D1	3948	-0,02959	0,01453	0,05860	0,51863
D2	3942	0,08035	0,14204	0,20264	0,00001
D3	3930	0,12460	0,21060	0,29345	0,00000
D4	3906	0,14811	0,26822	0,38053	0,00002
D5	3858	0,13730	0,30688	0,45898	0,00056
D6	3762	0,06892	0,31527	0,52538	0,01333
D7	3570	-0,01262	0,34924	0,63022	0,05854
D8	3186	0,17843	0,63285	0,86477	0,01134
S8	3186	0,68951	0,88809	0,96248	0,00001

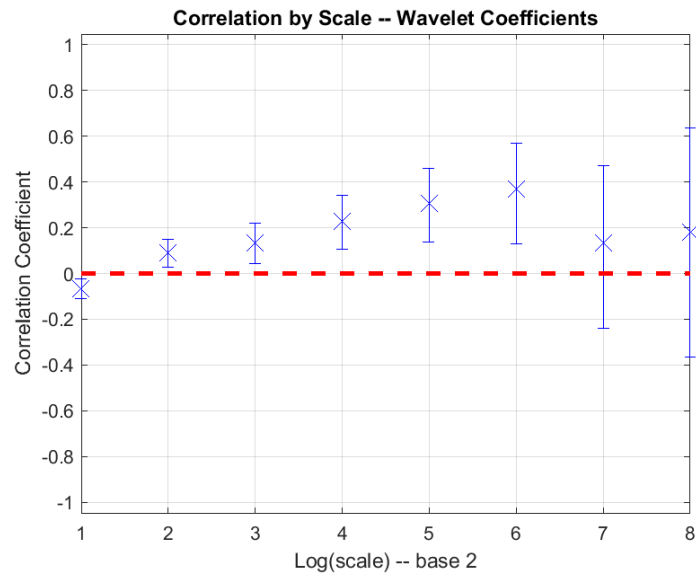


Figura 4.21: Gráfico da correlação multiescala dos preços de fechamento do Bitcoin e *Dow Jones* de 2012 até 2023

### 4.2.3 Mapa de coerência

Como visto no gráfico 4.10 algumas séries como o do *Dow Jones* e do BOVESPA apresentam momentos com correlação positiva e em outros momentos correlação negativa, por exemplo Janeiro de 2020 e Outubro de 2022 respectivamente. Por isso que no gráfico 4.22 temos que as séries do *Dow Jones* e IBOVESPA são correlacionadas em alguns momentos em relação à tendência e em outros em relação à frequência. Em relação a tendência, as séries apresentam alta magnitude quadrática de coerência a partir de 2016 onde a série do IBOVESPA também apresenta uma tendência de crescimento. Para níveis medianos e com frequência mais alta, as séries se correlacionam fortemente quando ocorre a crise causada pelo COVID-19. Assim fica comprovado com a análise gráfica de coerência

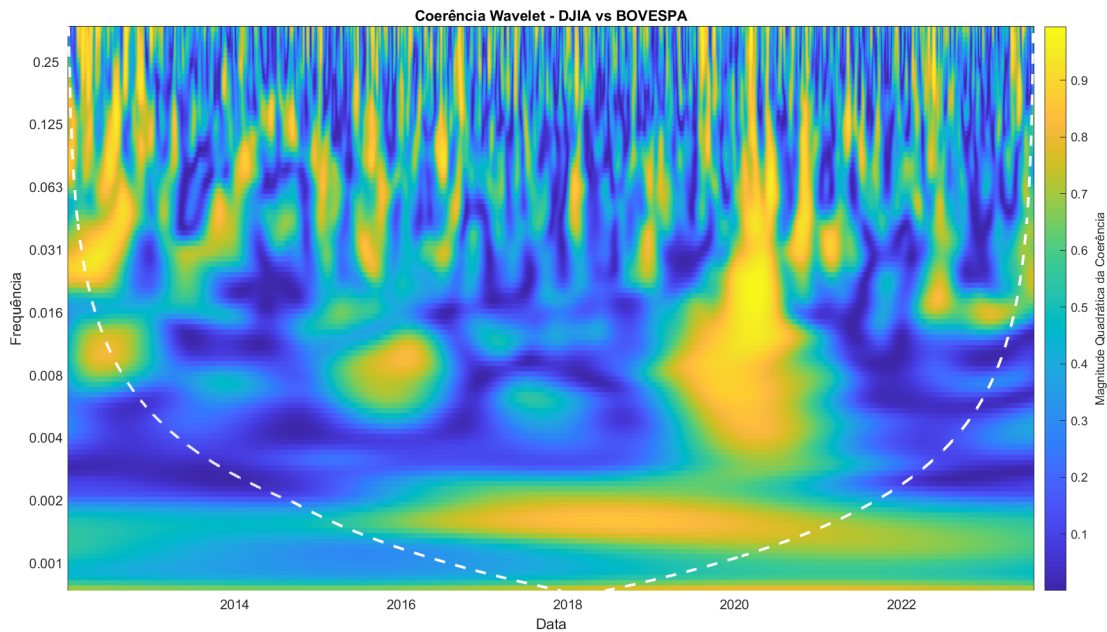


Figura 4.22: Gráfico do mapa de coerência dos preços de fechamento entre o IBOVESPA e *Dow Jones* de 2012 até 2023

*wavelet* que essas séries são correlacionadas.

Se nota no gráfico 4.23 que ao longo dos níveis das decomposições temos que as séries do *Dow Jones* e S&P500 são altamente correlacionadas. Porque o gráfico apresenta uma magnitude quadrática de coerência próxima de 1, ou amarela. Isto se deve, porque as séries são de índices que têm as mesmas empresas e também que o comportamento do sinal são parecidos. Assim fica comprovado com a análise gráfica de coerência *wavelet* que essas séries são fortemente correlacionadas.

No gráfico 4.24 que ao longo dos níveis das decomposições temos que as séries do *Dow Jones* e Bitcoin são correlacionadas apenas em relação a tendência. Porque o gráfico apresenta uma magnitude quadrática de coerência próxima de 1, ou amarela, em níveis em que a série apresenta baixa frequência e alta amplitude. Mas em níveis com alta frequência e baixa amplitude a magnitude apresenta uma fraca correlação. Isto se dá porque o crescimento que foi identificado depois de 2020 ocorre com o Bitcoin e o *Dow Jones*.

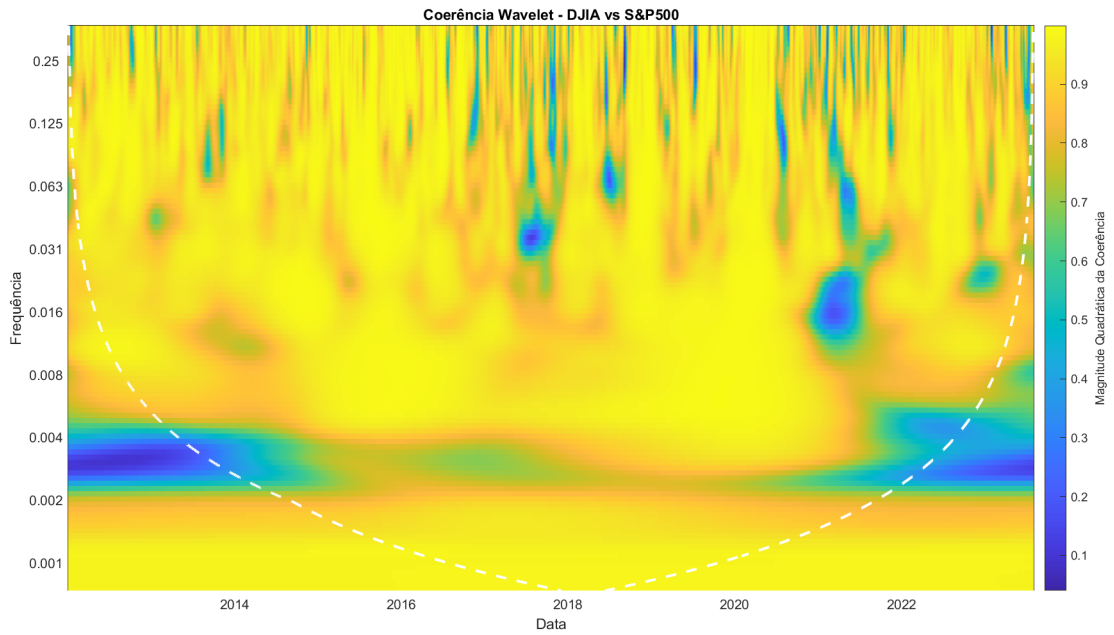


Figura 4.23: Gráfico do mapa de coerência dos preços de fechamento entre o S&P500 e *Dow Jones* de 2012 até 2023

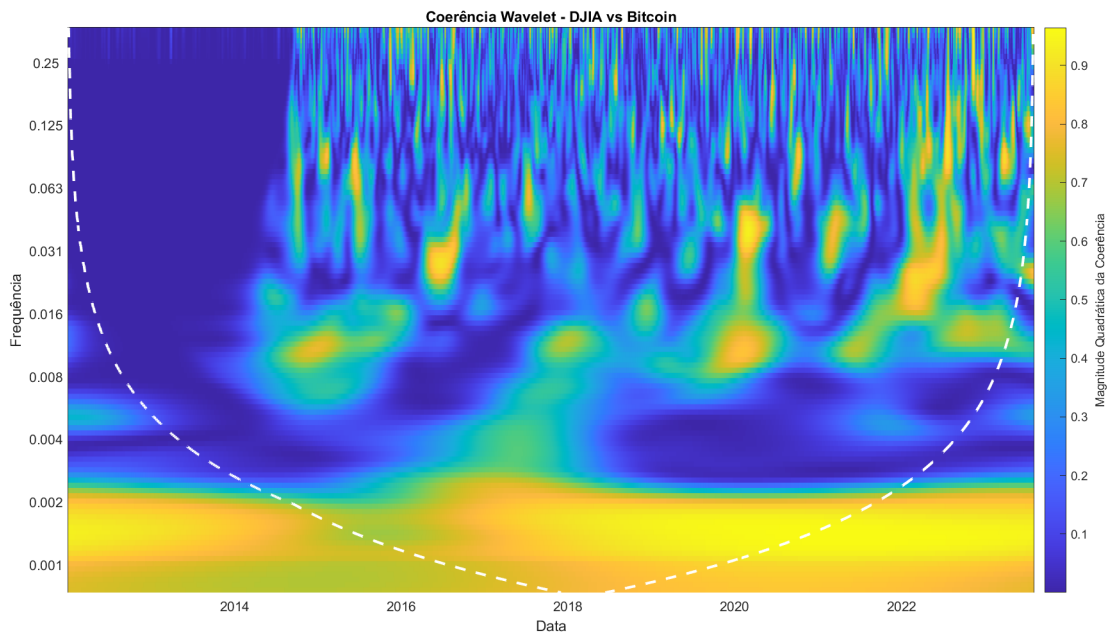


Figura 4.24: Gráfico do mapa de coerência dos preços de fechamento entre o Bitcoin e *Dow Jones* de 2012 até 2023



# Capítulo 5

## Conclusão

Em nosso estudo, escolhemos sem nenhum critério algumas séries financeiras para avaliarmos a correlação entre elas através de técnicas via *wavelet*. Onde aplicamos um método chamado de interpolação para substituir dados faltantes do banco de dados que foi construído através de uma API. Logo após uma análise descritiva e identificamos que principalmente a tendência como um fato comum entre os sinais de estudo.

Quando partimos para a transformada *wavelet* ficou mais claro que a crise causada pelo COVID-19 era um período influente em todas as séries em estudo. E ao decompor a série percebemos que a suposição criada na análise descritiva de que a tendência era algo parecido em todos os sinais. Além disso, confirmamos tal suposição pelas correlações multiescalas, que apresentaram correlações fortes e significativas para os níveis de tendência entre as séries.

Através deste estudo de transformações de *wavelet* fui capaz de ampliar meus conhecimentos na área das séries temporais, que é pouco estimulada no curso de graduação de estatística, mas muito utilizada por economistas e engenheiros.

Sobre os aprendizados na área da programação, fui estimulado a utilizar o MATLAB, que é um software pouco utilizado pelo curso. E Python para gerar gráficos em três dimensões para transformações *wavelet*.

Como a soma das decomposições resulta no sinal original, então para um estudo futuro será interessante aplicar modelos preditivos diferentes para cada nível de decomposição com a finalidade de obter uma predição mais realista. Onde o foco seria em modelar a variância nos primeiros níveis e depois modelar a média em níveis maiores. Outra oportunidade é ajustar um modelo para uma série financeira, utilizando níveis de decomposição que são significativamente correlacionados para se obter um melhor ajuste aos modelos.



# Referências Bibliográficas

- Aguiar-Contraria, L., S. M. (2011). Oil and the macroeconomy: using wavelets to analyze old issues. *Empir Econ* 40, página 645–655.
- Barbosa, A. . e Blitzkow, D. . (2008). *Ondaletas : Histórico e Aplicação*.. Instituto de Astronomia, Geofísica e Ciências Atmosféricas da Universidade de São Paulo – IAG/USP; (2) Escola Politécnica da Universidade de São Paulo – EPUSP – PTR – LTG.
- Daubechies, I.; Mallat, S. W. A. (1992). Introduction to the special issue on wavelet transforms and multiresolution signal analysis. *IEEE Trans. Inf. Theory*, **38**, n. **2**, 528–531.
- GRINSTED, R.; JEVREJEVA, S. M. J. I. (2003). Influência da arctic oscilação e el nino-oscilação sul (enos) no gelo condições no mar báltico: a abordagem wavelets. *J. Geophys. Res.*, **108**, n. **21**., 4677.
- GURGEL, V. (2013). *Aplicação de Técnicas Waveletts em análise de séries temporais para detecção de correlações*.. Mossoró: Universidade Federal Rural do Semi-Árido.
- Gurley, K., T. M. e Kareem, A. (1997). Analysis and simulation tools for wind engineering. *Probab. Eng. Mech.*, 12(1), página 9–31.
- HYNDMAN, R. J. e ATHANASOPOULOS, G. (2018). *Forecasting: principles and practice*.. OTexts.
- K. Gurley, M.ASCE; T. Kijewski, S. e A. Kareem, M. (2002). First- and higher-order correlation detection using wavelet transforms. *Journal of Engineering Mechanics, Issue 2 (February 2003)*, **129**.
- MORETTIN, P. A. e TOLOI, C. M. C. (2018). *Análise de séries temporais: modelos lineares univariados*.. Editora Blucher.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.  
*<https://bitcoin.org/bitcoin.pdf>*.

Percival, Donald B & Walden, A. T. (2000). *Wavelet Methods for Time Series Analysis*.  
Cambridge University Press.

WOODWARD, W. A., G. H. L. e. E. A. C. (2017). *Applied time series analysis with R*.  
CRC press.



# Apêndice A

## Apêndice

### Códigos do MatLab

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%% Decomposicao de uma serie %%%%%%%%%%%
x = dados.SeP500;
t = dados.Date;
mra = modwtmra(modwt(x,8));
helperMRAPlot(x,mra,t,'wavelet','Wavelet ARM - S&P500')

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%% Correlação entre duas séries %%%%%%%%%%%
x = dados.bitcoin;
y = dados.DJIA;
[rho,pval] = corrcoef(x,y);
wtPI = modwt(x,'db1',8,'reflection');
wtGE = modwt(y,'db1',8,'reflection');
wcorrtable = modwtcorr(wtPI,wtGE,'db2',0.95,'reflection','table');
display(wcorrtable)

piwt = modwt(x,'fk8',8);
pcwt = modwt(y,'fk8',8);
figure;
modwtcorr(piwt,pcwt,'fk8')
```

```

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Mapa de coerência entre duas séries %
t = dados.Date;
x = dados.DJIA;
y = dados.bitcoin;
[wcoh,~,period,coi] = wcoherence(x,y);

figure
h = pcolor(t,log2(period),wcoh);
h.EdgeColor = 'none';
ax = gca;
ytick=round(pow2(ax.YTick),3);
ax.YTickLabel=ytick;
ax.XLabel.String='Data';
ax.YLabel.String='Frequência';
ax.Title.String = 'Coerência Wavelet - DJIA vs Bitcoin';
hcol = colorbar;
hcol.Label.String = 'Magnitude Quadrática da Coerência';
hold on;
plot(ax,t,log2(coi),'w--','linewidth',2)

```

### Códigos em Python

```

# Instalando a biblioteca para a coleta dos dados
!pip install yfinance --upgrade --no-cache-dir

# Importante bibliotecas
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import pandas_datareader.data as web
import datetime
import missingno as msno

```

```

import pywt
import yfinance as yf
yf.pdr_override()

# Obtendo os dados do mercado
ibov = web.get_data_yahoo('^BVSP')
dji = web.get_data_yahoo('^DJI')
SeP500 = web.get_data_yahoo("^GSPC")
bitcoin = web.get_data_yahoo("BTC-USD")

# Criando um dataframe com os dados de fechamento
df_ = pd.concat([ibov["Close"], dji["Close"], SeP500["Close"], bitcoin["Open"]],
                axis=1, join='outer')
df_.columns = ["IBOV", "DJIA", "SeP500", "bitcoin"]
df_o = df_['2012-01-01': '2023-08-04']
df_o.tail()

# Visualize os valores ausentes como uma matriz
msno.matrix(df_o[["IBOV", "DJIA", "bitcoin", "SeP500"]])
msno.bar(df_o[["IBOV", "DJIA", "bitcoin", "SeP500"]])

ax = df_o.IBOV["2022-01-01": "2023-01-01"].plot(figsize=(18,10), color='C62402')
ax.set_xlabel("Data", fontsize=20)
ax.set_ylabel("Preço de fechamento (em reais)", fontsize=20)
plt.title('Preço de fechamento do BOVESPA de 2022 até 2023', fontsize=30)
plt.legend(loc='upper left', fontsize=20)
plt.tight_layout()
plt.style.use('seaborn-bright')
plt.grid(True)
plt.show()

# Substituindo valores faltantes
df_o.IBOV = df_o.IBOV.interpolate(method="backfill")

```

```

df_o.DJIA = df_o.DJIA.interpolate(method="backfill")
df_o.SeP500 = df_o.SeP500.interpolate(method="backfill")
df_o.bitcoin = df_o.bitcoin.interpolate(method="backfill")

# Cópia do DF
df = df_o.copy()

# Visualizacao de dados

## Teste de estacionariedade
from statsmodels.tsa.stattools import adfuller
def test_stationarity(timeseries, NOME):

    #Determining rolling statistics
    rolmean = timeseries.rolling(window=30).mean()
    rolstd = timeseries.rolling(window=30).std()

    #Plot rolling statistics:
    orig = plt.plot(timeseries, color='blue',label=NOME)
    mean = plt.plot(rolmean, color='red', label='Média movel')
    std = plt.plot(rolstd, color='black', label = 'Var movel')
    plt.legend(loc='best')
    plt.title('Média e variância moveis')
    plt.show(block=False)

    #Perform Dickey-Fuller test:
    print("Resultado do teste de Dickey-Fuller:")
    dftest = adfuller(timeseries, autolag='AIC')
    dfoutput = pd.Series(dftest[0:4], index=['Test Statistic', 'p-value',
                                           '#Lags Used', 'Number of Observations Used'])
    for key,value in dftest[4].items():
        dfoutput['Critical Value (%s)'%key] = value
    print(dfoutput)

```

```
## Visualizacao das series financeiras
### IBOVESPA
ax = df[["IBOV"]].plot(figsize=(18,10), color='C62402')
ax.set_xlabel("Data", fontsize=20)
ax.set_ylabel("Preço de fechamento (em reais)", fontsize=20)
plt.title('Preço de fechamento do Índice IBOVESPA', fontsize=30)
plt.legend(loc='upper left', fontsize=20)
plt.tight_layout()
plt.style.use('seaborn-bright')
plt.grid(True)
plt.show()

test_stationarity(df["IBOV"], "IBOVESPA")

### Dow Jones
ax = df[["DJIA"]].plot(figsize=(18,10), color='blue')
ax.set_xlabel("Data", fontsize=20)
ax.set_ylabel("Preço de fechamento (em dolares)", fontsize=20)
plt.title('Preço de fechamento do Dow Jones', fontsize=30)
plt.legend(loc='upper left', fontsize=20)
plt.tight_layout()
plt.style.use('seaborn-bright')
plt.grid(True)
plt.show()

test_stationarity(df["DJIA"], "DJIA")

### Bitcoin
ax = df.bitcoin.plot(figsize=(18,10), color='green')
ax.set_xlabel("Data", fontsize=20)
ax.set_ylabel("Preço de fechamento (em dolares)", fontsize=20)
plt.title('Preço de fechamento do Bitcoin', fontsize=30)
```

```
plt.legend(loc='upper left', fontsize=20)
plt.tight_layout()
plt.style.use('seaborn-bright')
plt.grid(True)
plt.show()
```

```
test_stationarity(df["bitcoin"], "Bitcoin")
```

```
### S&P500
```

```
ax = df.SeP500.plot(figsize=(18,10), color='black')
ax.set_xlabel("Data", fontsize=20)
ax.set_ylabel("Preço de fechamento (em dolares)", fontsize=20)
plt.title('Preço de fechamento do SP500', fontsize=30)
plt.legend(loc='upper left', fontsize=20)
plt.tight_layout()
plt.style.use('seaborn-bright')
plt.grid(True)
plt.show()
```

```
test_stationarity(df["SeP500"], "S&P500")
```

```
### Todas as series juntas
```

```
ax = df["DJIA"].plot(figsize=(18,10), color='blue')
df["IBOV"].plot(color = "C62402")
df["bitcoin"].plot(color='green')
df["SeP500"].plot(color='black')
ax.set_xlabel("Data", fontsize=20)
ax.set_ylabel("Preço de fechamento", fontsize=20)
plt.title('Preços de fechamento do Dow Jones, BOVESPA, S&P500 e Bitcoin', fontsize=30)
plt.legend(loc='upper left', fontsize=20)
plt.tight_layout()
plt.style.use('seaborn-bright')
plt.grid(True)
```

```

plt.show()

# Transformacao Wavelet e escalograma
### IBOVESPA
serie_d = df["DJIA"].copy()

amostras_pad = 200
serie_pad = pywt.pad(serie_d, amostras_pad, 'symmetric')

####Obtém o ano inicial e final
ano_inicial = min(serie_d.index).year
ano_final = max(serie_d.index).year

####Cria indices para mostrar no eixo x
anos = np.arange(ano_inicial, ano_final + 1, 1)
ia = np.arange(0, len(serie_d), 330)

####Calcula o escalograma
wav = 'gaus1' # Wavelet com 1 momento nulo
s = np.arange(1, 100) #Escalas utilizadas
coef, freqs = pywt.cwt(serie_pad, s, wav, method='fft') ####coeficientes wavelets
coef_DJIA = np.abs(coef)**2 #obtém o Escalograma

####Trunca W devido a operação de inserção de amostras (padding)
coef_DJIA = coef_DJIA[:, amostras_pad:-amostras_pad]

plt.figure(dpi=150, figsize=(10, 4))
plt.subplot(121)
plt.plot(serie_d)
plt.ylabel('Dow Jones Index - Preço de fechamento')
plt.xlabel('Ano')
plt.grid()

```

```

plt.subplot(122)
plt.imshow(coef_DJIA, extent=[0, len(serie_d), s[-1], s[0]], cmap='jet',
           aspect='auto', interpolation='spline16')
plt.xticks(ia, anos, rotation='vertical')
plt.xlabel('Translação ( $\tau$ )')
plt.ylabel('Escala ( $s$ )')
plt.colorbar()
plt.style.use('fast')
plt.grid(False)
plt.tight_layout()
plt.show()

```

```
# Visualização 3D
```

```

from mpl_toolkits.mplot3d import Axes3D
import matplotlib.pyplot as plt
import pandas as pd
import seaborn as sns

```

```
coef_df = pd.DataFrame(coef_DJIA)
```

```

df_3d=coef_df.unstack().reset_index()
df_3d.columns=["X","Y","Z"]
df_3d['X']=pd.Categorical(df_3d['X'])
df_3d['X']=df_3d['X'].cat.codes

```

```
datas = pd.Series(df.index.to_list())
```

```
datas_ = []
```

```
for j in range(len(datas)):
```

```
    for i in range(0,99):
```

```
        datas_.append(datas[j])
```

```
df_3d["Date"] = datas_
```



```
fig = plt.figure(figsize = (15,16))
ax = fig.add_subplot(projection = '3d')
ax.plot_trisurf(df_3d['Y'], df_3d['X'], df_3d['Z'], cmap="jet", linewidth=0.2)
ax.set_xlabel('$Escala$', fontsize=15)
ax.set_ylabel('$t$', fontsize=15)
ax.set_zlabel('$Coefientes TWC$', fontsize=15, rotation = 0)
plt.style.use('fast')
ax.view_init(30,10)
```