



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2020/05.21.13.54-TDI

**METODOLOGIA NUMÉRICO-COMPUTACIONAL  
PARA ANÁLISE MULTIESCALA DE SINAIS:  
APLICAÇÕES A PROCESSOS ELETRODINÂMICOS  
ESPACIAIS**

Rodrigo Takeshi Seo

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pela Dra. Margarete Oliveira Domingues, aprovada em 20 de maio de 2020.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/42H34DB>>

INPE  
São José dos Campos  
2020

**PUBLICADO POR:**

Instituto Nacional de Pesquisas Espaciais - INPE  
Gabinete do Diretor (GBDIR)  
Serviço de Informação e Documentação (SESID)  
CEP 12.227-010  
São José dos Campos - SP - Brasil  
Tel.:(012) 3208-6923/7348  
E-mail: pubtc@inpe.br

**CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE - CEPPII (PORTARIA Nº 176/2018/SEI-INPE):****Presidente:**

Dra. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos Climáticos (CGCPT)

**Membros:**

Dra. Carina Barros Mello - Coordenação de Laboratórios Associados (COCTE)

Dr. Alisson Dal Lago - Coordenação-Geral de Ciências Espaciais e Atmosféricas (CGCEA)

Dr. Evandro Albiach Branco - Centro de Ciência do Sistema Terrestre (COCST)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia e Tecnologia Espacial (CGETE)

Dr. Hermann Johann Heinrich Kux - Coordenação-Geral de Observação da Terra (CGOBT)

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação - (CPG)

Silvia Castro Marcelino - Serviço de Informação e Documentação (SESID)

**BIBLIOTECA DIGITAL:**

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Serviço de Informação e Documentação (SESID)

**REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:**

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação (SESID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SESID)

**EDITORAÇÃO ELETRÔNICA:**

Ivone Martins - Serviço de Informação e Documentação (SESID)

Cauê Silva Fróes - Serviço de Informação e Documentação (SESID)



MINISTÉRIO DA  
CIÊNCIA, TECNOLOGIA  
E INOVAÇÕES



sid.inpe.br/mtc-m21c/2020/05.21.13.54-TDI

**METODOLOGIA NUMÉRICO-COMPUTACIONAL  
PARA ANÁLISE MULTIESCALA DE SINAIS:  
APLICAÇÕES A PROCESSOS ELETRODINÂMICOS  
ESPACIAIS**

Rodrigo Takeshi Seo

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pela Dra. Margarete Oliveira Domingues, aprovada em 20 de maio de 2020.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/42H34DB>>

INPE  
São José dos Campos  
2020

Dados Internacionais de Catalogação na Publicação (CIP)

---

Seo, Rodrigo Takeshi.

Se61m Metodologia numérico-computacional para análise multiescala de sinais: aplicações a processos eletrodinâmicos espaciais / Rodrigo Takeshi Seo. – São José dos Campos : INPE, 2020.

xxviii + 128 p. ; (sid.inpe.br/mtc-m21c/2020/05.21.13.54-TDI)

Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2020.

Orientadora : Dra. Margarete Oliveira Domingues.

1. Eletrodinâmica espacial. 2. Magneto-hidrodinâmica.  
3. Clima espacial. 4. Análise wavelet. 5. Análise de similaridade.  
I.Título.

CDU 517.44:629.78

---



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÕES  
**INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS**

## **FOLHA DE APROVAÇÃO**

A FOLHA DE APROVAÇÃO SERÁ INCLUIDA APÓS RESTABELECIMENTO DAS ATIVIDADES PRESENCIAIS.

Por conta da Pandemia do COVID-19, as defesas de Teses e Dissertações são realizadas por vídeo conferência, o que vem acarretando um atraso no recebimento nas folhas de aprovação.

Este trabalho foi aprovado pela Banca e possui as declarações dos orientadores (confirmando as inclusões sugeridas pela Banca) e da Biblioteca (confirmando as correções de normalização).

Assim que a Biblioteca receber a Folha de aprovação assinada, esta folha será substituída.

Qualquer dúvida, entrar em contato pelo email: [pubtc@inpe.br](mailto:pubtc@inpe.br).

Divisão de Biblioteca (DIBIB).



*“Meninos mimados não podem reger a nação”.*

KLEBER CAVALCANTE GOMES  
em *“Espiral de ilusão”*, 2017



*a mulher de braços fortes, minha mãe*



## AGRADECIMENTOS

Agradeço ao Instituto Nacional de Pesquisas Espaciais (INPE) e ao Curso de Pós-graduação em Computação Aplicada (CAP) a infraestrutura necessária ao desenvolvimento deste trabalho.

Agradeço à CAPES o suporte financeiro sem o qual não seria possível a realização deste trabalho. Deixo também um grande obrigado a todos os cidadãos brasileiros que valorizam e prestigiam a ciência brasileira.

Agradeço à minha orientadora Dra. Margarete Oliveira Domingues o suporte profissional e pessoal tão necessários, ao longo de todo esse período, para essa realização.

Agradeço também aos professores da CAP todos os conhecimentos obtidos.

Aos membros da proposta, em especial, ao professor Dr. Odim Mendes as valiosas sugestões e inúmeras discussões científicas.

Agradeço à minha avó Etsuko Massuda Seo todo o carinho e apoio em todos os momentos de minha vida, minhas tias Norma e Nanci e especialmente minha falecida mãe Laura, sem a luta dessas mulheres nada disso seria possível.

Agradeço à minha companheira Rebeca Prado toda paciência e amor durante todo o mestrado.

Agradeço aos amigos Anna Karina, Muller Lopes e José Marchezi a minha inspiração em coleguismo e amizade para que eu conseguisse dar este grande passo.

Agradeço aos companheiros Vinícius Deggeroni, Raphael Moreno e todos os colegas da CAP os momentos de estudos e de descontração.

Agradeço às minhas amigas e fisioterapeutas Naiara Oliveira e Amanda Denari que me auxiliaram e me recuperaram em um dos momentos mais difíceis da minha vida.

Agradeço à Faculdade de Tecnologia (FATEC) de São José dos Campos e a todos os professores e colegas do curso de Tecnologia em Banco de Dados, que me forneceram as bases sólidas de conhecimento.

Agradeço aos colegas da *Quero Educação*, em especial ao time de Data Science e Quero Pago, que me deram toda a disponibilidade e convivência necessárias para a finalização deste trabalho.



## RESUMO

A computação científica constitui uma área de desenvolvimento de novos conceitos e ferramentas úteis a muitas áreas de atuação, sobretudo atendendo as necessidades crescentes da área das investigações e de aplicações das ciências espaciais, como do emergente tópico de clima espacial. Com a disponibilidade de avançadas metodologias de medição, o monitoramento do ambiente solar-terrestre tem gerado grandes quantidades de dados com alta resolução temporal e maior extensão espacial, contendo fenômenos das mais diversas escalas espaço-temporais, o que demanda assim abordagens multiescalas para a análise desses fenômenos. Aliado a isso com a melhoria dos recursos computacionais e métodos numéricos, a simulação das equações magneto-hidrodinâmicas está se tornando essencial na análises dos fenômenos eletrodinâmicos considerando uma visão em volume extenso do espaço e da evolução dos fenômenos. Esta dissertação apresenta o desenvolvimento de uma metodologia numérico-computacional de análise multiescala da comparação de medidas e de resultados de modelagem magneto-hidrodinâmica de processos de plasma espacial, como caso de estudo os dados da Sonda Cassini na atmosfera de Titan, satélite de Saturno. Utilizando diversas técnicas *wavelet* em conjunto com outras ferramentas de medidas pode-se avaliar quantitativamente a similaridade dos sinais. Complementarmente, estudou-se também casos de efeito de perturbação magnética, devido ao acoplamento eletrodinâmico, avaliado na superfície terrestre para caracterização de sinais empregando-se a metodologia proposta. A importância deste trabalho vem de se dispor de uma metodologia quantitativa que auxilie no entendimento desses processos físicos e de como seus comportamentos dinâmicos estão sendo representados.

Palavras-chave: Eletrodinâmica Espacial. Magneto-hidrodinâmica. Clima Espacial. Análise *Wavelet*. Análise de Similaridade. Análise de sinais.



# COMPUTATIONAL NUMERICAL METHODOLOGY FOR MULTISCALE SIGNAL ANALYSIS: APPLICATIONS TO SPATIAL ELECTRODYNAMIC PROCESSES

## ABSTRACT

Scientific computing is an area of development of new concepts and tools for many areas of expertise, above all meeting needs of growing areas like space science research and applications, such as the emerging topic of space weather. With the availability of advanced measurement methodologies, solar-terrestrial environment monitoring has generated a significant amount of data with spatial-temporal high resolution, containing phenomena with many space-temporal scales, demanding multiscale approach for these phenomena analysis. Allied to this with the improvement of computational resources and numerical methods, the simulation of magnetohydrodynamics equations is becoming essential in detecting electrical phenomena considering the broad view and space volume evolution of phenomena. This work presents the development of a numerical-computational methodology of multiscale analysis of the comparison of measures and magneto-hydrodynamic modelling of space plasma processes, such as data from the Cassini Probe in the atmosphere of Titan, Saturn's natural satellite. Obtained from several wavelet techniques in conjunction with other measurement tools, the similarity evaluation of the signals is quantitatively performed. Complementarily, cases of disturbance effects of electrodynamic coupling assessed on the Earth's surface are studied using the proposed methodology to characterize signals. The importance of this work comes from having a robust quantitative methodology that helps in understanding the physical processes and how their dynamic behaviours could be represented.

Keywords: Space Electrodynamics. Magnetohydrodynamics. Space weather. Wavelet analysis. Similarity analysis. Signal analysis.



## LISTA DE FIGURAS

	<u>Pág.</u>
1.1 Representação do plasma do vento solar atingindo o campo magnético terrestre e produzindo interações. . . . .	2
2.1 Esquema de interação dos dados. . . . .	9
2.2 Domínio e elementos básicos de um modelo MHD . . . . .	12
2.3 Exemplo de campo magnético medido (MAG) pela sonda Cassini em passagem próxima a Titan, satélite de Saturno, em 26 de dezembro de 2005, intervalo em horas pelo tempo universal da Terra. MHD e HALL indicam respectivamente os resultados da simulação magneto-hidrodinâmica resistiva e com o termo Hall. . . . .	13
2.4 Dados do vento solar e do meio interplanetário - janeiro/2018. Dados obtidos de OMNI web service, NASA. . . . .	14
3.1 Variação dos parâmetros $\tau$ e $a$ de uma função <i>wavelet</i> . . . . .	18
3.2 Plotagem da variação do parâmetro $\omega_0$ da <i>wavelet</i> de Morlet . . . . .	19
3.3 Valor de integração da função <i>wavelet</i> de Morlet . . . . .	20
3.4 Série temporal sintética e seus respectivos escalogramas relativos a diferentes famílias de Morlet. . . . .	22
3.5 Fase dos coeficientes <i>wavelet</i> . . . . .	23
3.6 Espectro <i>wavelet</i> global (em negrito) e espectro de Fourier. . . . .	24
3.7 Série temporal de manchas solas solares, escalograma, fase e espectro. . . . .	25
3.8 Série temporal $f_2$ , escalograma, espectros e fase. . . . .	26
3.9 Série temporal original e série temporal derivada. . . . .	27
3.10 Série temporal derivada de $f_2$ , escalograma, espectros e fase. . . . .	28
3.11 Série Temporal $f_2$ Original e Filtrada. . . . .	29
3.12 Escalograma, Fase e Espectro da série $f_2$ filtrada . . . . .	29
3.13 Entropia <i>wavelet</i> da série temporal $f_1$ . . . . .	31
3.14 Entropia <i>Wavelet</i> das série de manchas solares. . . . .	31
3.15 Séries temporais $B_z$ e AE . . . . .	34
3.16 Escalograma de $B_z$ e de AE, Coerência <i>wavelet</i> e diferença de fase . . . . .	35
3.17 Correlação por escalas de $B_z$ e AE módulo e fase. . . . .	36
4.1 Esquema da DWT decimada e ortogonal . . . . .	40
4.2 Esquema da MODWT . . . . .	42
4.3 Decomposição multi-nível da componente $B_z$ . . . . .	47
4.4 Série Temporal $B_z$ , Filtragens com a DWT e MODWT . . . . .	48

5.1	Comparativo entre Distância Euclidiana e DTW . . . . .	50
5.2	Caminho $\mathfrak{W}$ . . . . .	52
5.3	Cenário 1: Séries Temporais, Matriz de Distância e Caminho $\mathfrak{W}$ . . . . .	53
5.4	Cenário 2: Séries Temporais, Matriz de Distância e Caminho $\mathfrak{W}$ . . . . .	54
5.5	Cenário 3: Séries Temporais, Matriz de Distância e Caminho $\mathfrak{W}$ . . . . .	55
5.6	Cenário 4: Séries Temporais, Matriz de Distância e Caminho $\mathfrak{W}$ . . . . .	56
5.7	Decomposição em escalas - $B_z$ . . . . .	58
5.8	Aplicação: Séries Temporais e Matriz de Distância. . . . .	60
5.9	Caminho $\mathfrak{W}$ do campo magnético . . . . .	61
6.1	Componente $B_z$ do campo magnético medido pela sonda Cassini . . . . .	64
6.2	Dados da componente $B_z$ : escalogramas e fases. . . . .	65
6.3	Espectro <i>wavelet</i> global para componente $B_z$ . . . . .	67
6.4	Módulo e fase da correlação por escala da componente $B_z$ . . . . .	68
6.5	Módulo e fase da correlação por escala dos dados da componente $B_z$ dos modelos rMHD e hMHD. . . . .	69
6.6	Entropia <i>wavelet</i> da componente $B_z$ e dos valores dos modelos MHD. . . . .	69
6.7	Coerência <i>wavelet</i> e diferença de fases para a relações medida $B_z$ e dado rMHD, $B_z$ e dado hMHD e entre rMHD e hMHD. . . . .	71
6.8	Sinais reconstruídos, pela aplicação da técnica <i>wavelet</i> , para os dados das medidas e das simulações. . . . .	72
6.9	Comparativo de matrizes de distância e caminhos $\mathfrak{W}$ entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita). . . . .	74
6.10	Escalogramas e fases do período calmo dos Dados dos índices AE e ASY-H. . . . .	77
6.11	Espectro global <i>wavelet</i> para AE e ASY-H, no período calmo. . . . .	78
6.12	Período calmo AE e ASY-H: correlação por escalas e fase. . . . .	79
6.13	Coerência <i>wavelet</i> e diferença de fase para AE e ASY-H no período calmo. . . . .	80
6.14	Entropia <i>wavelet</i> dos índices AE e ASY-H no período calmo. . . . .	80
6.15	Sinais reconstruídos por filtragem <i>wavelet</i> para os índices AE e ASY-H no período calmo. . . . .	81
6.16	Matriz de Distância e Caminho $\mathfrak{W}$ para AE e ASY-H, no período calmo. . . . .	82
6.17	Escalograma e fases dos índices AE e ASY-H em período perturbado. . . . .	83
6.18	Espectro global <i>wavelet</i> para AE e ASY-H, no período perturbado. . . . .	84
6.19	Correlação por escalas e fase para AE e ASY-H no período perturbado. . . . .	85
6.20	Coerência <i>wavelet</i> e diferença de fase para AE e ASY-H, no período perturbado. . . . .	85
6.21	Entropia <i>wavelet</i> para AE e ASY-H, período perturbado. . . . .	86
6.22	Sinais reconstruídos pela técnica <i>wavelet</i> para AE e ASY-H. . . . .	86
6.23	Matriz de Distância e Caminho $\mathfrak{W}$ para AE e ASY-H, para período perturbado. . . . .	87

A.1	Filtros de suavização. . . . .	99
A.2	Exemplo de coerência <i>wavelet</i> com janela <i>Box Car</i> . . . . .	100
A.3	Exemplo de coerência <i>wavelet</i> com janela <i>Gaussiana</i> . . . . .	102
B.1	Dados do módulo do campo magnético $B$ , escalogramas e fases. . . . .	105
B.2	Espectro <i>wavelet</i> global para o módulo do campo magnético $B$ referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado). . . . .	106
B.3	Módulo e fase da correlação por escala de $B$ em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul, pontilhado) e hMHD (vermelho, tracejado). . . . .	107
B.4	Módulo e fase da correlação por escala dos dados modelados pelo rMHD e hMHD. . . . .	108
B.5	Dados de $B$ , coerência e diferença de fase. . . . .	109
B.6	Dados de $B$ - Entropia Wavelet. . . . .	110
B.7	Comparativo de matrizes de distância e caminhos $\mathfrak{W}$ entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita). . . . .	111
B.8	Dados da componente $B_x$ , escalogramas e fases. . . . .	112
B.9	Espectro <i>wavelet</i> global para componente $B_x$ referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado). . . . .	113
B.10	Módulo e fase da correlação por escala da componente $B_x$ em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul, pontilhado) e hMHD (vermelho, tracejado). . . . .	114
B.11	Módulo e fase da correlação por escala dos dados modelados pelo rMHD e hMHD. . . . .	114
B.12	Dados da componente $B_x$ , coerência e diferença de fase. . . . .	115
B.13	Componente $B_x$ - Entropia Wavelet. . . . .	116
B.14	Comparativo de matrizes de distância e caminhos $\mathfrak{W}$ entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita). . . . .	118
B.15	Dados da componente $B_y$ , escalogramas e fases. . . . .	120
B.16	Espectro <i>wavelet</i> global para componente $B_y$ referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado). . . . .	121
B.17	Módulo e fase da correlação por escala da componente $B_y$ em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul, pontilhado) e hMHD (vermelho, tracejado). . . . .	122
B.18	Módulo e fase da correlação por escala dos dados modelados pelo rMHD e hMHD. . . . .	123

B.19	Dados da componente $B_y$ , coerência e diferença de fase. . . . .	124
B.20	Componente $B_y$ - Entropia <i>wavelet</i> . . . . .	125
B.21	Comparativo de matrizes de distância e caminhos $\mathfrak{W}$ entre dados obser- vacionais e dos modelos rMHD (a esquerda) e hMHD (a direita). . . . .	127

## LISTA DE TABELAS

	<u>Pág.</u>
1.1 Aspectos metodológicos e tipos de resultados no contexto da dissertação. . . . .	5
4.1 Valores não-zero dos filtros passa-baixa de Daubechies. . . . .	46
5.1 Comparativo de distâncias Euclidiana e DTW. . . . .	57
5.2 Composição de escalas das séries temporais. . . . .	59
5.3 Comparativo de distâncias Euclidiana e DTW. . . . .	62
6.1 Similaridades entre os dados observacionais e dos modelos rMHD e hMHD referentes à componente do campo magnético $B_z$ . . . . .	75
6.2 Similaridades entre os dos índices geomagnéticos AE e ASY-H no periodo geomagneticamente perturbado. . . . .	88
B.1 Similaridades entre os dados observacionais e dos modelos rMHD e hMHD referentes ao campo magnético. . . . .	128



## LISTA DE ABREVIATURAS E SIGLAS

AE	–	Índice Auroral Electroject
asyH	–	Índice Asymmetric H
CWT	–	Continuous Wavelet Transform
DB8	–	Daubechies 8
DOG	–	Derivated of Gaussian
DWT	–	Discrete Wavelet Transform
DST	–	Storm Time Disturbance
DTW	–	Dynamic Time Warping
EDP	–	Equações Diferenciais Parciais
FFT	–	Fast Fourier Transform
HILDCAA	–	High Intensity, Long Duration, Continuous AE Activity
hMHD	–	Modelo Magneto-Hidrodinâmico de Hall
MHD	–	Magneto-Hidrodinâmica
MODWT	–	Maximum-overlap Discrete Wavelet Transform
MR	–	Multiresolução
rMHD	–	Modelo Magneto-Hidrodinâmico Resistivo
RMSE	–	Root Mean Square Error
SYM-H	–	Longitudinally symmetric horizontal component
WISA	–	Wavelet-based index of magnetic storm activity
XWT	–	Cross Wavelet Transform



## LISTA DE SÍMBOLOS

$t$	– tempo físico
$\mathbf{B}$	– campo magnético interplanetário
$ \mathbf{B} $	– magnitude do campo magnético interplanetário
$B_x$	– componente $x$ do campo magnético interplanetário
$B_y$	– componente $y$ do campo magnético interplanetário
$B_z$	– componente $z$ do campo magnético interplanetário
$\rho$	– densidade do plasma
$p$	– pressão do plasma
$\gamma$	– índice adiabático
$\nabla \cdot$	– operador de divergência
$\mathbf{u}$	– vetor velocidade
$\mathbf{J}$	– vetor densidade corrente
$\mathcal{E}$	– densidade total de energia
$C_\psi$	– condição de admissibilidade
$\psi$	– função <i>wavelet</i> analisadora
$\omega_0$	– frequência central associada a uma função <i>wavelet</i>
$a$	– escala
$\tau$	– deslocamento temporal
$W_\psi^f$	– coeficientes <i>wavelet</i>
$\theta$	– fase dos coeficientes <i>wavelet</i>
$S_\psi^f$	– espectro <i>wavelet</i> global
$W_\psi^{f,g}$	– transformada <i>wavelet</i> cruzada de duas funções $f$ e $g$ quaisquer
$R^{f,g}$	– coerência <i>wavelet</i>
$\phi^{f,g}$	– fase da coerência <i>wavelet</i>
$C_{S_\psi}$	– correlação por escalas
$d_k^j$	– detalhes da DWT
$j$	– níveis de decomposição da DWT
$\mathcal{D}$	– distância Euclidiana
$\mathcal{Q}$	– distância DTW
$\mathfrak{M}$	– matriz de distâncias DTW
$\mathfrak{W}$	– <i>warping path</i>



## SUMÁRIO

	<u>Pág.</u>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
<b>2 EMBASAMENTO FÍSICO E DADOS</b> . . . . .	<b>7</b>
2.1 O cenário eletrodinâmico espacial . . . . .	7
2.2 Dados . . . . .	8
2.3 Visão conceitual de modelo MHD gerador de dados . . . . .	10
2.4 Dados selecionados para análise . . . . .	12
2.4.1 Medidas e resultados de simulação . . . . .	12
2.4.2 Medidas do meio interplanetário próximo à Terra . . . . .	13
2.4.3 Dados classificadores de processos físicos no ambiente terrestre . . . . .	15
<b>3 ANÁLISE POR TRANSFORMADAS <i>WAVELET</i> CONTÍNUAS</b> <b>17</b>	
3.1 Funções <i>wavelet</i> analisadoras . . . . .	17
3.1.1 Wavelet de Morlet . . . . .	18
3.2 Transformada wavelet contínua . . . . .	20
3.2.1 Ferramentas usuais da CWT . . . . .	21
3.2.2 Análise e resultados em aplicações conhecidas . . . . .	24
3.2.3 Outras ferramentas da transformada <i>wavelet</i> contínua . . . . .	30
3.2.3.1 Entropia <i>wavelet</i> . . . . .	30
3.2.3.2 Transformada <i>Wavelet</i> Cruzada - XWT . . . . .	32
3.2.3.3 Coerência <i>Wavelet</i> . . . . .	32
3.2.3.4 Correlação por escala . . . . .	33
3.2.4 Análises preliminares das ferramentas de estudo de relações tempo- escala entre séries temporais . . . . .	33
3.2.4.1 Resultados parciais: relações entre dados do meio interplanetário e de superfície . . . . .	34
3.2.5 Implementação da CWT e ferramentas . . . . .	36
<b>4 ANÁLISE POR TRANSFORMADAS <i>WAVELET</i> DISCRETAS</b> <b>37</b>	
4.1 Definição de uma Análise Multirresolução (MR) . . . . .	37
4.2 Algoritmos . . . . .	38
4.2.1 Comentários sobre os algoritmos e códigos utilizados . . . . .	44
4.2.2 Resultados dos testes de sensibilidade . . . . .	46

<b>5</b>	<b>ANÁLISE POR MÉTRICAS DE SIMILARIDADE . . . . .</b>	<b>49</b>
5.1	Distância Euclidiana . . . . .	49
5.2	Dynamic Time Warping - DTW . . . . .	50
5.3	Testes de sensibilidade . . . . .	51
5.3.1	Estudo sobre o <i>Warping Path</i> . . . . .	52
5.3.2	Cenário 1: diferentes frequências . . . . .	52
5.3.3	Cenário 2: diferentes fases . . . . .	54
5.3.4	Cenário 3: deslocamento temporal . . . . .	55
5.3.5	Cenário 4: localização de eventos . . . . .	55
5.3.6	Resumo e comparação dos cenários . . . . .	57
5.4	Resultados parciais: estudo de uma componente do campo magnético interplanetário . . . . .	57
<b>6</b>	<b>APLICAÇÕES, RESULTADOS E DISCUSSÕES . . . . .</b>	<b>63</b>
6.1	Aplicação da metodologia a dados observacionais de satélite e dados de simulação . . . . .	63
6.1.1	Análise de relações tempo-escala . . . . .	64
6.1.2	Análise das métricas de similaridade entre dados observacionais e dados simulados . . . . .	73
6.2	Aplicação da metodologia a dados classificadores de distúrbios geomag- néticos . . . . .	76
6.2.1	Análise dos dados para período geomagneticamente calmo . . . . .	76
6.2.1.1	Análise das métricas de similaridade entre dados dos índices geomag- nético no período geomagneticamente calmo . . . . .	81
6.2.2	Análise dos dados em período geomagneticamente perturbado . . . . .	83
6.2.2.1	Análise de métricas de similaridade entre dados dos índices geomag- néticos no período geomagneticamente perturbado . . . . .	87
<b>7</b>	<b>CONCLUSÃO . . . . .</b>	<b>91</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>93</b>
<b>A</b>	<b>Apêndice A: Estudos sobre os efeitos de suavização da coerência <i>wavelet</i> . . . . .</b>	<b>99</b>
<b>B</b>	<b>Apêndice B: Análise complementar dos Dados da Sonda Cassini</b>	<b>105</b>
B.1	Módulo do campo magnético . . . . .	105
B.1.1	Análise das métricas de similaridade entre dados observacionais e dados simulados . . . . .	110

B.2	Componente magnética na direção x . . . . .	111
B.2.1	Análise das métricas de similaridade entre dados observacionais e dados simulados . . . . .	117
B.3	Componente magnética na direção y . . . . .	119
B.3.1	Análise das métricas de similaridade entre dados observacionais e dados simulados . . . . .	126
B.4	Resumo dos valores das métricas de similaridade . . . . .	127



# 1 INTRODUÇÃO

Cada vez mais, de forma perceptível e sob fácil justificativa, constata-se que a computação científica vem contribuindo e participando de forma notável nos desenvolvimentos científico, tecnológico, industrial e de serviços e, por consequência, no progresso da sociedade. Isso se dá, sobretudo, por auxiliar na criação de novos conceitos e ferramentas em todos esses setores. Nestas primeiras décadas do Século XXI, uma das áreas principais em que se pode apontar um crescimento significativo é no contexto das Ciências Espaciais.

Por necessidades mesmo indissociáveis desta área, verificam-se as utilizações de técnicas e de recursos computacionais na vanguarda de novos descobrimentos, ainda mais com a demanda de conceitos que contribuem para a exploração e análise de grandes volumes de dados e, em especial, na capacidade de lidar mesmo com o desafio da complexidade desses dados.

Um dos recursos poderosos da computação científica, que pode contribuir significativamente nas análises estruturadas de sinais, particularmente pela representação ou simplificação destes, é o oriundo da Matemática Aplicada em que se faz o seu desenvolvimento com fundamentação e aplicação da análise multiescala, que, por sua vez, serve-se de alguns conceitos fundamentais da física.

Para contextualizar e embasar o campo da aplicação prática do trabalho aqui desenvolvido, e que atende aos esforços de missão do INPE, apresenta-se a seguir, em uma visão sucinta, a fenomenologia espacial sob interesse.

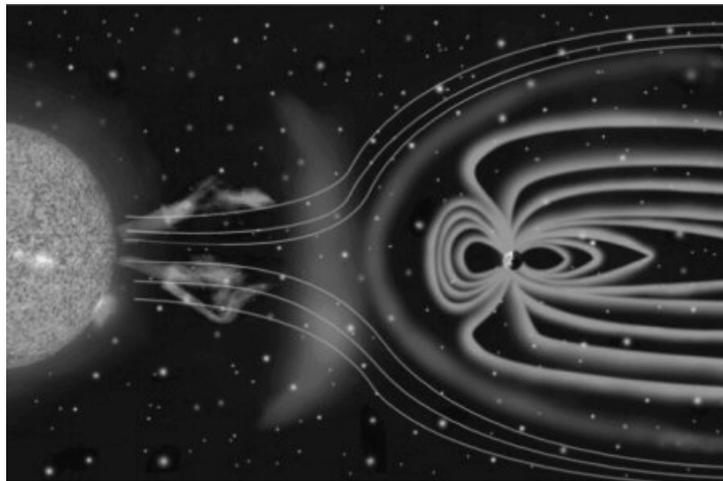
O Sol, como estrela mais próxima, afeta o ambiente em seu entorno e, por consequência, a Terra e os artefatos espaciais humanos. O monitoramento continuado desse, assim designado, ambiente espacial fornece uma grande quantidade de dados científicos, contendo implicitamente informações sobre as variabilidades dos eventos eletrodinâmicos espaciais, em especial, no espaço próximo (KAMIDE; CHIAN, 2007; MOLDWIN, 2008; MENDES, 1992).

Tais informações podem ser utilizadas em modelos que descrevem os fenômenos decorrentes do acoplamento Sol-Terra devido à interação do vento solar com a magnetosfera terrestre (OGINO et al., 1986).

O vento solar é um fluxo de plasma – composição de partículas eletricamente carregadas que cumprem certas propriedades coletivas de comportamento – que se destaca do Sol devido a diferença de pressão entre a coroa solar e o meio interplane-

tário, carrega consigo linhas do campo magnético solar, no que é denominado campo magnético do meio interplanetário (MOLDWIN, 2008; KIVELSON; RUSSELL, 1995). A Figura 1.1 é uma representação artístico-conceitual do vento solar agindo sobre o campo magnético terrestre, que tem uma alteração da morfologia, e produzindo diversos processos eletrodinâmicos nessa atmosfera terrestre.

Figura 1.1 - Representação do plasma do vento solar atingindo o campo magnético terrestre e produzindo interações.



Fonte: Adaptado de ESA (2018).

Essa interação cria um entorno da Terra muito rico de fenômenos, em particular os eletrodinâmicos, que ocorrem em diversas escalas de espaço e de tempo, que, para pesquisas mais aprofundadas, requerem a consideração de abordagem do tipo multiescala (MENDES et al., 2005b).

Ilustrando com uma aplicação cuja abordagem é pertinente a este trabalho, esse conceito matemático-numérico é utilizado, por exemplo, no trabalho de Lorenzo et al. (2016), possibilitando simplificar em tempo-escala os dados de entrada de uma modelagem magneto-hidrodinâmica (MHD), com o propósito de suavizar as variações e tornar tratável a entrada de valores a serem utilizados na ferramenta de simulação numérica.

Por um outro lado, como um sentido prático da relevância dessa abordagem, a geração ou o estabelecimento de alertas considerando os impactos de perturbações eletrodinâmicas, com resultados como as perturbações geomagnéticas, são viabiliza-

dos pela disponibilidade ou existência de técnicas de medição, avaliação e previsão dos comportamentos em sinais, ou seja, nas séries temporais (SHUE et al., 1998; MEJNERTSEN et al., 2018). No domínio de sinais ou de processos físicos, há ainda uma outra característica adicional, que se mostra necessária nas análises, que é a capacidade de avaliar a proximidade ou consistência da informação gerada por modelos confrontada com os dados reais. Para isso, busca-se estruturar então abordagem técnica que calcule e estabeleça uma avaliação de proximidade de informações.

Um dos grandes desafios primordiais nas atividades científicas é lidar tanto com medições, caracterizando-as objetivamente, quanto estabelecer uma avaliação de consistência entre as informações envolvidas, fatos que estabelecem a motivação para os esforços deste trabalho. Dentro do âmbito dos estudos espaciais, que visa também aplicações potenciais para programas de clima espacial, as ferramentas da matemática aplicada e a computação possuem grande participação na estruturação de conceitos matemático-numérico-computacionais, criação de metodologias físico-matemático-computacionais, e elaboração de tecnologias inovadoras voltadas a aplicações, focadas aqui na possibilidade de auxiliar quando da execução de simulações e da investigação de experimentos de grande escala temporal e espacial.

Como visão prática, o esforço da análise de resultados computacionais realizada de forma sistemática, assertiva e ágil torna-se necessário para que seja possível a geração de informações confiáveis e precisas, neste caso, referente aos estudos sobre possíveis impactos eletrodinâmicos provenientes das interações Sol-Terra. Neste trabalho, define-se como objetivo principal a criação de uma metodologia de análise envolvendo elementos da matemática aplicada para multirresolução e a implementação de recursos numérico-computacionais integrados, possibilitando uso conjunto de várias técnicas, para essa análise. O que se mostra como significativa e original contribuição deste trabalho, uma vez que se pode dispor de uma metodologia de análise multiescala e de abordagens que auxiliem na caracterização dos sinais, explicitando as assinaturas de processos e eventuais estruturas físicas implícitas, e a determinação de similaridade das informações.

Desta forma, enquanto bases da metodologia, enfocam-se técnicas de estudo de sinais que envolvessem recursos de tratamento de aspectos globais e locais. Para possibilitar o estabelecimento de uma análise integrada de sinais, estudam-se em particular a Transformada *Wavelet* Contínua (CWT, Continuous *Wavelet* Transform), a Transformada *Wavelet* Discreta (DWT, Discrete *Wavelet* Transform), considerando também ferramentas derivadas respectivamente, e técnicas de similaridade como a da

distorção temporal dinâmica (DTW, *Dynamic Timing Warping*), que provê métricas de avaliação de distâncias temporais. Estabelecem-se e demonstram-se, igualmente, para cada técnica algumas de suas potenciais aplicações. A grande importância de uma metodologia construída com esses recursos é a sua utilização para uma investigação de dados oriundos de processos não lineares e regimes não estacionários, como os são os fenômenos da física espacial.

Como base para a aplicação e teste de conceito da metodologia integrada desenvolvida, utilizam-se parâmetros medidos por satélites e resultados modelados por código MHD da região espacial do entorno de um astro e, ainda complementarmente, medidas relativas ao acoplamento eletrodinâmico do plasma do vento solar e da magnetosfera terrestre por meio de quantificações de perturbações geomagnéticas obtidas na superfície. A abordagem metodológica busca ainda identificar e mensurar a similaridade das grandezas físicas envolvidas nos processos físicos, o que deve propiciar um entendimento mais fidedigno com a realidade. A Tabela 1.1 apresenta a compilação da identificação do aspecto da metodologia de análise e a informação alcançada com o resultado obtido.

Tabela 1.1 - Aspectos metodológicos e tipos de resultados no contexto da dissertação.

Ítem	Análise	
	Aspecto	Resultado
<b>Transf. <i>wavelet</i> contínua</b>		
1	Escalograma (módulo)	Contribuição da energia do sinal em relação as escalas ao longo do tempo
2	Escalograma (fase)	a fase do sinal ao longo do tempo-escala
3	Espectro global (integrado)	Contribuição de energia do sinal por escalas, integrado no tempo
4	Entropia <i>wavelet</i> (integrado)	Nível de informação de um sinal integrado nas escalas
<b>Bivariada</b>		
5	Transformada <i>wavelet</i> cruzada	Produto ponto a ponto entre os escalogramas de 2 sinais
6	Coerência <i>wavelet</i>	Avaliação de correlação entre escalas no tempo entre 2 sinais
	Fase (coerência)	Avaliação de entrosamento de fase entre 2 sinais
7	Correlação por escala (módulo)	Avaliação da correlação por escala integrado no tempo entre 2 sinais
	Correlação por escala (fase)	Avaliação da fase da correlação por escala integrado no tempo entre 2 sinais
<b>Transf. <i>wavelet</i> discreta</b>		
8	Decomposição ortogonal	Decomposição em níveis (escala) de sinais e recurso de filtragem, redução de dados entre os níveis
9	Decomposição redundante	Decomposição em níveis (escala) de sinais e recurso de filtragem, mantém número de dados entre escalas
<b>Métricas de similaridade</b>		
10	Distorção dinâmica do tempo	Avaliação da alteração no tempo da relação entre sinais

Fonte: Produção do Autor.

O trabalho está organizado, portanto, da forma seguinte. No Capítulo 2, abordam-se o contexto do ambiente físico, as medições e quantidades envolvidas na avaliação de processos espaciais de interesse, e elementos de modelagem numérica, considerando

simulações magneto-hidrodinâmicas, para investigações de tais processos. Detalham-se também os dados utilizados no trabalho e o âmbito das suas escolhas. Por último, evidenciam-se as características de uma metodologia de análise de dados a serem atendidas. Nos Capítulos 3, 4 e 5 apresentam-se, desta forma, a estruturação da metodologia de análise multiescala, em uma visão integradora, com o uso de ferramentas ou técnicas selecionadas considerando transformadas *wavelets* e análise de similaridade em séries temporais. Em cada um desses três capítulos são estudadas, exemplificadas e/ou testadas as técnicas em casos típicos. No Capítulo 6, aplica-se a metodologia desenvolvida e apresentam-se os resultados alcançados, com a sua interpretação e discussão no contexto da computação científica. No Capítulo 7, resumam-se as conclusões do trabalho. Para completeza, os apêndices trazem informações complementares ou acessórias que ajudam o entendimento e revelam a extensão do trabalho realizado.

## 2 EMBASAMENTO FÍSICO E DADOS

Neste capítulo, as bases para a compreensão do ambiente espacial e dos processos de eletrodinâmica são consideradas de forma abrangente no conteúdo mas com apresentação concisa, orientadas para o objetivo do trabalho.

Os dados e as quantidades do meio espacial são descritos e também relacionados aos processos. As caracterizações para o entendimento dos dados, que possibilitam a estruturação de informações, são adiante propostas, o que dá base para uma metodologia de análise, em especial de análise integrada dos dados.

### 2.1 O cenário eletrodinâmico espacial

O Sol, como uma estrela, caracteriza-se por produzir energia por reação nuclear em sua parte central decorrente de compressão devido à atração gravitacional. Essa energia mantém um entorno constituído de partículas eletricamente carregadas que, por obedecerem certas propriedades coletivas, compõem um ambiente de plasma. Esse plasma porta um campo magnético estelar que se estende das regiões mais interiores para as camadas externas e, devido a um processo de expansão desse plasma, posteriormente é transportado para regiões do espaço mais distantes da estrela. Três processos principais são agentes de atuação eletrodinâmica do Sol sobre o espaço e os corpos à sua volta. Radiação eletromagnética é emitida e ao incidir tanto energiza quanto ioniza as atmosferas planetárias. Ainda que de forma mais localizada, partículas energéticas de energia produzem também esses efeitos. Por último, manifestações de plasma solar magnetizado, em seu deslocamento pelo meio interplanetário, podem interagir com um astro em sua trajetória.

De um ponto de vista geral, a atmosfera ionizada e um campo magnético existente de um planeta sob a atuação do plasma solar magnetizado possibilitam estabelecer uma região eletrodinâmica, designado de magnetosfera, de exclusão do plasma solar em torno do planeta, porém preenchido de um plasma planetário (KIVELSON; RUSSELL, 1995). A Figura 1.1, anteriormente vista, destaca, do seu centro para a direita, essa situação de acoplamento eletrodinâmico, que se revela de muita importância em especial para o cotidiano na Terra.

Fundamentalmente, a magnetosfera é a região da atmosfera permeada do campo magnético que controla o movimento das partículas (MENDES, 1992; MENDES et al., 2005). A dinâmica de interação entre as partículas provenientes do Sol e a magnetosfera terrestre ocorre quando o vento solar comprime o campo magnético no

lado diurno, formando uma onda de choque envoltória (*bow shock*). Ao passar pela onda de choque, o vento solar perde velocidade e é desviado ao redor do astro, para uma região denominada bainha magnética (*magnetosheath*), e na parte posterior forma uma cauda magnética (*magnetotail*). A magnetopausa (*magnetopause*) é a camada que separa o plasma do vento solar do plasma predominantemente planetário (SCHUNK; NAGY, 2009). A extensão da magnetosfera está diretamente relacionado com a pressão dinâmica do vento solar, enquanto a orientação e intensidade do campo magnético interplanetário define a taxa do fluxo de energia e transferência de partículas. Toda essa dinâmica de acoplamento Sol-astro é origem de fenômenos das mais variadas ordens de escala temporal e espacial, produzindo efeitos eletrodinâmicos que, para exemplo, no caso da Terra, impactam diretamente dispositivos eletrônicos e toda a sociedade (MENDES et al., 2005).

Portanto, o acoplamento do plasma solar com a atmosfera planetária inicia uma interação eletrodinâmica, cuja intensidade de efeitos é modulada por características dos agentes solares. Há no ambiente da magnetosfera do astro uma ionização atmosférica, a geração de plasmas, a ocorrência de ondas eletromagnéticas, o surgimento de campos e correntes elétricas e, ainda, perturbações magnéticas (KIRCHOFF, 1991).

A Física Espacial possibilitou nos tempos atuais estabelecer uma área de aplicação prática das Ciências Espaciais, designada como Clima Espacial. Para os propósitos de melhor monitoramento e mesmo uma futura capacidade de previsão nesta área emergente, inúmeros estudos, desenvolvimentos de metodologias e implementações de ferramentas inovadoras revelam-se extremamente necessários. Este trabalho busca contribuir com esses esforços de melhor atender os requisitos para investigações mais amplas, quanto à abrangência espacial, mais intensivas, quanto ao volume de dados e continuidade de medições, e melhor caracterizadoras, quanto à diversidade de processos e comportamentos multiescalas. Com este fim, este trabalho faz uso de medições dentro da magnetosfera, dos resultados obtidos de modelagens magneto-hidrodinâmicas referente a tais medições e, para um caso exemplo, de quantidades mensuradas na avaliação de efeitos de perturbações magnéticas.

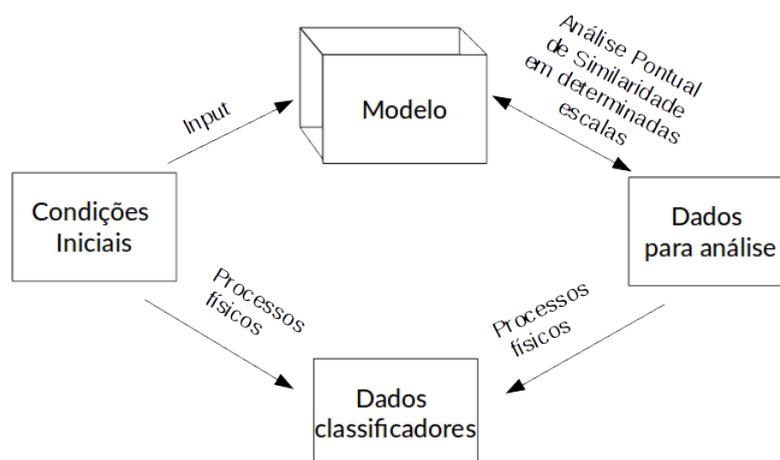
## 2.2 Dados

Os dados utilizados neste trabalho possuem origens diversas e são subdivididos em dois grupos básicos. Em primeiro lugar, os dados que compõem as condições iniciais e de fronteira de modelo magneto-hidrodinâmico (*input*) consistem em medidas do meio interplanetário coletados fora da influência do astro em estudo (que não estarão sendo considerados aqui) e os dados tanto das medições por sonda espa-

cial quanto da simulação numérico-computacional respectiva para o entorno de um astro (o que será considerado neste trabalho). Basicamente, o resultado ou output da simulação se dá no domínio de uma caixa tridimensional envolvendo a região de interesse. Em segundo lugar, os dados de satélite localizados no meio interplanetário nas proximidades da Terra mas fora da magnetosfera (não considerados aqui) e os dados classificadores da influência no ambiente terrestre, representados por índices geomagnéticos (analisados neste trabalho). A complexidade do estudo proposto, no tempo disponível de um mestrado, implicou a escolha de um conjunto limitado mas representativo de dados, que caracterizassem quantificações relacionadas, e possibilitassem constituir efetivamente um estudo de conceito de desenvolvimento de metodologia numérico-computacional por técnica de análise multiescala de sinais e técnica de similaridade em séries temporais, em abordagem integradora.

A idéia central do desafio e esforço do trabalho é buscar validar e relacionar os dados envolvidos em modelagens computacionais e também os dados de resposta a parâmetro interplanetário que se caracterizam como seus efeitos na atmosfera planetária. A Figura 2.1 apresenta um esquema de como esses conjuntos de dados, que são uma parcela de um grande conjunto de dados possíveis de se considerar no futuro, interagem neste trabalho.

Figura 2.1 - Esquema de interação dos dados.



Fonte: Produção do Autor.

Embora possam se aplicar em um contexto geral, essa representação de interação de dados é totalmente consistente com as condições de investigação dos efeitos ao

ambiente espacial, por exemplo, sobre a Terra. Com medições constituindo as condições iniciais, pode-se ter as entradas para uma ferramenta de modelagem como para estruturação de quantidades classificadoras. Tanto da modelagem alcançam-se quantidades representativas de processos a serem confrontados com medições, constituindo dados para análise, como dos classificadores também alcançam-se relações com os dados de análise, oriundos de processos de acoplamento eletrodinâmico. Uma metodologia que seja integradora das formas de análise mostra-se muito útil e necessária para a compreensão específica dos fenômenos como igualmente um auxílio à percepção do cenário total envolvido.

### 2.3 Visão conceitual de modelo MHD gerador de dados

Campos magnéticos influenciam e são influenciados por diversos fluxos de fluidos eletronicamente condutores observados na natureza. O estudo dessas relações de um ponto de vista macroscópico é chamado de abordagem Magneto-hidrodinâmica (MHD), em que se consideram observáveis (i.e. medições) macroscópicos. Os efeitos das interações descritas na MHD descrevem diversos processos reais, como, por exemplo, na indústria em métodos de solda ou fusão nuclear; na Astrofísica em processos de formação de estrelas no universo; na Heliofísica em interações entre o plasma solar e um campo geomagnético, que é de particular interesse neste trabalho. O modelo magneto-hidrodinâmico de interesse neste estudo é descrito por Equações Diferenciais Parciais (EDP), possibilitando simular a Equação de Euler para fluidos combinada com as Equações de Maxwell para o eletromagnetismo, como descrito em vários livros e trabalhos clássicos de plasma, como por exemplo [Bittencourt \(2013\)](#), [Davidson \(2001\)](#), [Goedbloed and Poedts \(2004\)](#), [Roberts \(1967\)](#), que, de forma específica, combinam as equações de Euler para fluidos com as equações de Maxwell para o eletromagnetismo.

O modelo MHD descreve a dinâmica do plasma como um fluido magnetizado utilizando o conjunto de variáveis macroscópicas  $(\rho, \rho\mathbf{u}, \mathcal{E}, \mathbf{B})$ , que correspondem às quantidades físicas de densidade, momento linear, energia interna e fluxo magnético, respectivamente.

Matematicamente, o modelo é descrito pelas equações:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho\mathbf{u}) = 0$$

$$\frac{\partial (\rho\mathbf{u})}{\partial t} + \nabla \cdot \left[ \rho\mathbf{u}\mathbf{u} + \left( p + \frac{\mathbf{B} \cdot \mathbf{B}}{2} \right) \mathbb{I} - \mathbf{B}\mathbf{B} \right] = \mathbf{0}$$

$$\frac{\partial \mathcal{E}}{\partial t} + \nabla \cdot \left[ \left( \mathcal{E} + p + \frac{\|\mathbf{B}\|^2}{2} \right) \mathbf{u} - (\mathbf{u} \cdot \mathbf{B}) \mathbf{B} + (\eta \mathbf{J}) \times \mathbf{B} + \mathbf{u}_H \left( \frac{\mathbf{B} \cdot \mathbf{B}}{2} \right) - \mathbf{B} (\mathbf{u}_H \cdot \mathbf{B}) \right] = 0$$

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \cdot \left[ \mathbf{u} \mathbf{B} - \mathbf{B} \mathbf{u} + \eta \left( (\nabla \mathbf{B})^T - \nabla \mathbf{B} \right) + \mathbf{u}_H \mathbf{B} - \mathbf{B} \mathbf{u}_H \right] = 0$$

em que  $\gamma$  é o índice adiabático,  $p$  é a pressão hidrodinâmica do plasma,  $\eta$  é a resistividade do plasma, a energia interna  $\mathcal{E}$  é dada pela combinação das energias hidrodinâmica e magnética e o termo de Hall é dado por  $\mathbf{u}_H = \frac{-\mathbf{J}}{ne}$ :

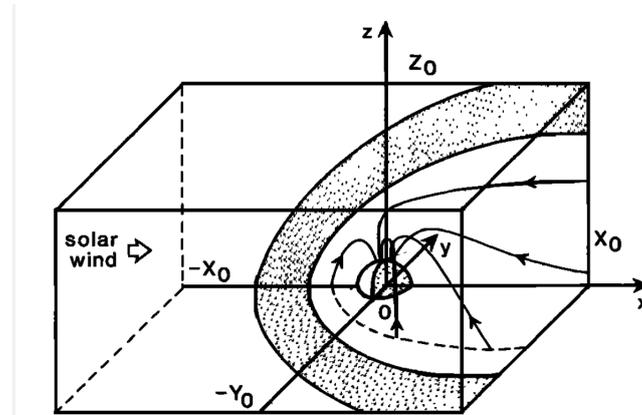
$$\mathcal{E} = \frac{p}{\gamma - 1} + \rho \frac{\mathbf{u} \cdot \mathbf{u}}{2} + \frac{\mathbf{B} \cdot \mathbf{B}}{2}. \quad (2.2)$$

A primeira equação do sistema MHD, chamada de equação da continuidade, descreve o transporte de matéria que ocorre durante a evolução temporal do plasma. A segunda equação (Equação do momento) é responsável por descrever a evolução do campo de velocidades do fluido. A terceira equação (Equação da energia) controla a evolução da energia interna do plasma ao longo de seu domínio. A última equação (Equação da indução) descreve a evolução do campo magnético durante a evolução do plasma. Para destacar as contribuições, utilizam-se a cor azul nos termos referentes à parte resistiva e a cor vermelha aos termos Hall.

Sob esta formulação, o modelo MHD ideal é caracterizado por  $\eta = 0$  e termo hall nulo, que permite ter os aspectos mais básicos da fluidodinâmica dos plasmas. O termo de resistividade associa-se a um efeito dissipativo sobre o campo magnético do plasma, o que permite a ocorrência de reconexões magnéticas (fusão de linhas de campo magnético em oposição de orientação). Por outro lado, ao utilizar o termo Hall, o modelo passa a ser capaz de representar um afastamento de centro de cargas portando cargas elétricas opostas e lidar com oscilações de alta frequência. Portanto, ao negligenciar a dissipação e o efeito hall, o modelo ideal conserva as quantidades de massa, momento, energia, helicidade, além de manter a topologia das linhas de campo magnético.

Os dados de entrada (condições iniciais e de fronteira) em modelagem MHD são obtidos das medições por satélites fora de um ambiente espacial que está sob o domínio magnético do astro em questão; enquanto que os de saída por meio da utilização dos resultados de simulação. Para exemplo, a Figura 2.2 traz o esquema do domínio para uma modelagem tridimensional MHD das interações do plasma solar com a magnetosfera terrestre. Detalhamentos, que não cabem no escopo deste trabalho, podem ser encontrados em [Ogino et al. \(1986\)](#).

Figura 2.2 - Esquemática do domínio e de elementos básicos de modelo magneto-hidrodinâmico em coordenadas cartesianas.



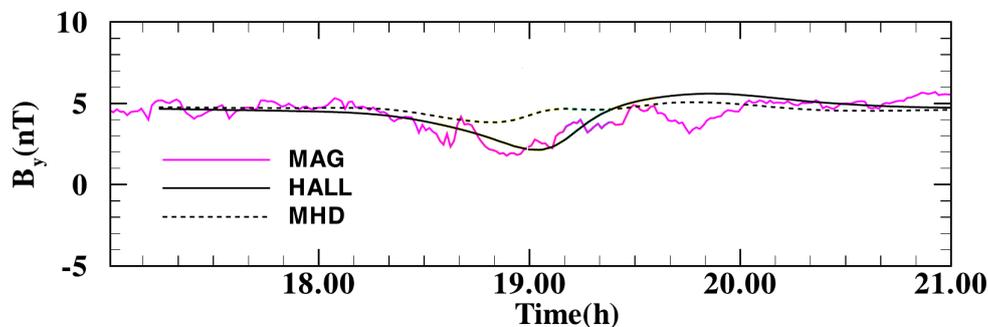
Fonte: Ogino et al. (1986).

## 2.4 Dados selecionados para análise

### 2.4.1 Medidas e resultados de simulação

Para análise de caso de modelagem magneto-hidrodinâmica, utilizam-se os dados da missão da sonda Cassini, NASA, em sua nona passagem em trânsito próximo à Titan, um satélite do planeta Saturno, em 26 de dezembro de 2005. Para informações mais detalhadas, pode-se consultar o trabalho de [Ma et al. \(2007\)](#). As medidas utilizadas para serem analisadas correspondem aos registros coletados pela sonda no intervalo de 17 a 21 horas, estabelecido em tempo universal da Terra. As variáveis consideradas são as componentes cartesianas do campo magnético  $B_x$ ,  $B_y$  e  $B_z$  (em nanoteslas), e a velocidade do plasma  $V$  (em quilômetros por segundo). A [Figura 2.3](#) mostra a medida do campo magnético  $B_y$ , para exemplo, com os respectivos resultados da simulação magneto-hidrodinâmica realizadas para a condição resistiva (MHD) e com o termo Hall (HALL). As demais figuras podem ser consultadas no próprio artigo mencionado. Para possibilitar o tratamento de análise, os dados foram digitalizados a partir das figuras de forma a recuperar os valores, permitindo aplicar a metodologia.

Figura 2.3 - Exemplo de campo magnético medido (MAG) pela sonda Cassini em passagem próxima a Titan, satélite de Saturno, em 26 de dezembro de 2005, intervalo em horas pelo tempo universal da Terra. MHD e HALL indicam respectivamente os resultados da simulação magneto-hidrodinâmica resistiva e com o termo Hall.



Adaptado de Ma et al. (2007).

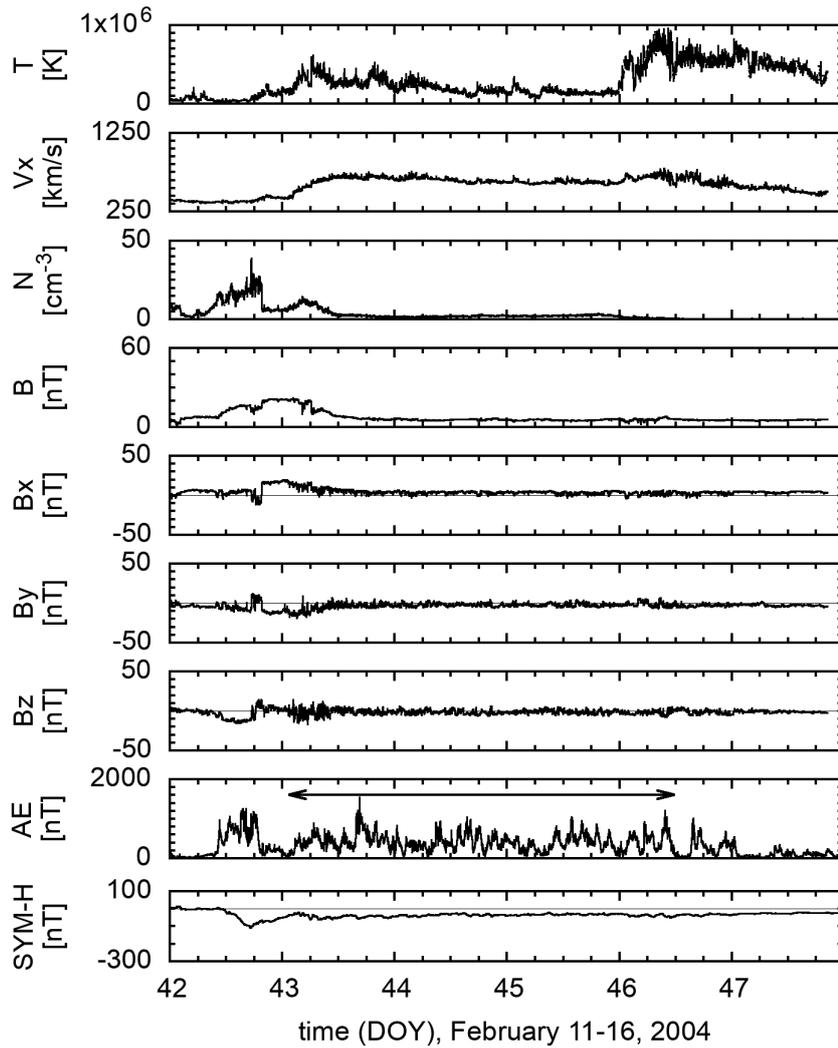
#### 2.4.2 Medidas do meio interplanetário próximo à Terra

De um ponto de vista geral, o campo magnético planetário age como um escudo protetor, devido à rigidez dielétrica, contra a radiação corpuscular solar, proporcionando um obstáculo efetivo para o deslocamento do plasma do vento solar (KIVELSON; RUSSELL, 1995). Isso define uma magnetosfera, região em que o campo magnético controla a dinâmica das partículas eletricamente carregadas (MENDES, 1992), conforme detalhamento já feito na Seção 2.1. Portanto, o acoplamento eletrodinâmico Sol-astro se inicia na interação do fluxo do plasma solar incidente sobre a magnetosfera de um astro (MENDES et al., 2005). Essa interação, por mecanismos diversos e complicados, vai originar perturbações magnéticas na atmosfera planetária, que podem, no caso da Terra, ser registradas na superfície.

No estudo desses acoplamentos, quanto a causas e efeitos, utilizam-se as medidas obtidas por satélites no meio interplanetário. Dados preparados com o auxílio de vários satélites estão disponíveis no portal OMNI web service, NASA, [https://omniweb.gsfc.nasa.gov/ow\\_min.html](https://omniweb.gsfc.nasa.gov/ow_min.html). Os dados de alta resolução temporal, cadência de um minuto de valor promediado, podem ser baixados do site [https://omniweb.gsfc.nasa.gov/form/omni\\_min.html](https://omniweb.gsfc.nasa.gov/form/omni_min.html). As medidas que são utilizadas nos estudos do acoplamentos consistem das componentes do campo magnético interplanetário ( $B_x$ ,  $B_y$  e  $B_z$ ), no sistema referencial do Sol-Terra-magnetosfera (Geocentric Solar Magnetospheric referencial System, GSM), em nanotestal, da den-

sidade de partículas  $N$  em partículas por centímetro cúbico, velocidade do plasma do vento solar  $V$ , e a temperatura do plasma  $T$  em kelvin. A Figura 2.4 apresenta, para exemplo, essas medidas para o intervalo de 11 a 16 de fevereiro de 2004, que correspondem em dias do ano (day of year, DOY) de 42 a 47.

Figura 2.4 - Dados do vento solar e do meio interplanetário - janeiro/2018. Dados obtidos de OMNI web service, NASA.



Fonte: Produção do autor.

A variável utilizada no trabalho será a componente do campo magnético interplanetário  $B_z$ , cuja orientação predominantemente norte ou sul, determina respectivamente interação viscosa nas fronteiras, relacionada a situação descrita como geo-

magneticamente calma, ou processo de reconexão magnética, relacionada a situação geomagneticamente perturbada. Como casos de estudos, dois eventos serão considerados no trabalho. O primeiro é o intervalo de 12 a 15 de fevereiro (DOY 43 a 46) de 2004, que refere-se à passagem de uma região de interação co-rotante (corrotating interaction region, CIR), que, pela manifestação de uma flutuação com orientação sul na medida  $Bz$ , produziu uma perturbação geomagnética fraca (com valor maior do que  $-50$  nT) em baixa latitude. O segundo é o intervalo de 18 a 21 de julho (DOY 199 a 202) de 2006, que refere-se a um período de variabilidade de baixa intensidade de  $Bz$ , redundando em um período geomagneticamente calmo.

### 2.4.3 Dados classificadores de processos físicos no ambiente terrestre

Disparadas primordialmente pelas flutuações com orientação sul do  $Bz$ , variações do campo magnético terrestre são produzidas em consequência e, desta forma, medidas por magnetômetros, que são aparelhos desenvolvidos e utilizados há muitos anos como sensores de campo magnético (KABATA; VITORELLO, 2007). Devido ao grande volume de dados gerados por esses aparelhos, tornando a sua manipulação complexa, índices geomagnéticos foram criados para caracterizar e classificar as variações e perturbações do campo geomagnético sem a necessidade de manipulação de volume massivo de dados (KIVELSON; RUSSELL, 1995).

O índice planetário Kp foi estabelecido como um indicador global de atividades do campo magnético terrestre. É obtido por meio de medições de treze observatórios geomagnéticos ao redor do globo e produz medições a cada três horas. Embora bastante útil, esse índice tem a limitação de não permitir a observação de variações de alta frequência devido ao seu intervalo de medições (CAMPBELL, 2003).

Introduzido originalmente em 1966 por Davis e Sugiura, o índice auroral AE caracteriza-se por quantificar as atividades geomagnéticas na região auroral do planeta. É obtido do processamento dos registros de doze observatórios geomagnéticos localizados nas zonas aurorais do hemisfério norte, produzindo apontamentos com amostragem em minutos (NOSE et al., 2011).

Outro índice comumente utilizado é o índice de perturbação geomagnética em médias para baixas latitudes relacionado a uma corrente elétrica anelar, em torno do equador, a aproximadamente seis raios terrestres, designado como Dst (YAMASHITA, 1999). Em função de variações para o sul e mais intensas em módulo de  $Bz$ , há um efeito de diminuição da componente horizontal do campo geomagnético em baixas latitudes. Isso cria o fenômeno de tempestades geomagnéticas, caracterizado por uma

fase de diminuição dessa componente atingindo um mínimo e depois com uma fase de recuperação até um valor de base ou de fundo. Esse índice quantifica esse comportamento (GONZALEZ et al., 1994). Uma limitação desse índice é que sua resolução é horário.

Um outro índice foi criado com as características do índice idealizado Dst, porém preparado para avaliar as contribuições de um corrente anelar simétrica, constituindo o índice SYM-H. Esse índice é calculado utilizando medições de 10 observatórios em médias para baixas latitudes (IYEMORI, 1990). A resolução temporal deste índice é de um minuto, conveniente às aplicações deste trabalho. Ao subtrair a componente simétrica anterior da resultante do campo magnético, possibilita-se definir um índice assimétrico, designado como ASY-H. O índice assimétrico possibilita quantificar um efeito de estruturas complicadas de correntes elétricas no entorno da Terra (CROOKER; SISCOE, 1971) e que aparentemente tem relação com deposições de energia na região auroral.

Os índices AE, SYM-H e ASY-H, utilizados neste trabalho, podem ser obtidos no portal OMNI web service, NASA, ([https://omniweb.gsfc.nasa.gov/form/omni\\_min.html](https://omniweb.gsfc.nasa.gov/form/omni_min.html)) ou no site do World Data Center for Geomagnetism, Kyoto (<http://wdc.kugi.kyoto-u.ac.jp/wdc/Sec3.html>). Também na Figura 2.4 estão representados, na parte inferior, esses índices geomagnéticos. O índice Dst possibilita constatar a ocorrência de uma tempestade geomagnética moderada (mínimo em aproximadamente  $-50$  nT) e o índice AE um perturbação auroral de valor moderado (acima de  $400$  nT. Sobre o AE, a flecha de dupla ponta indica o intervalo em que a perturbação caracterizou um fenômeno designado como Atividade Auroral Contínua de longa duração e alta Magnitude (HILDCAA, High Intensity Long Duration Auroral Continuous Activity).

A existência de parâmetros do ambiente solar-terrestre motiva e possibilita, assim, a construção de metodologias de análise integrada que auxiliem na interpretação dessa diversidade de fenômenos, por caracterização inequívoca dos sinais e, portanto, dos processos envolvidos.

### 3 ANÁLISE POR TRANSFORMADAS *WAVELET* CONTÍNUAS

Neste capítulo, como primeira parte para estabelecer uma metodologia integrada de análise de sinais, introduz-se e estuda a Transformada *Wavelet* Contínua (CWT, Continuous *Wavelet* Transform), considerando suas características. Em síntese, a CWT é uma ferramenta de análise exploratória dos sinais não estacionários no contexto tempo-escala. Ressalta-se ainda que esta apresenta redundância entre as escalas. Estabelecem-se e demonstram-se, igualmente, para a técnica algumas de suas potenciais aplicações.

Na sequência são apresentadas as propriedades necessárias para uma função ser considerada uma função *wavelet*.

#### 3.1 Funções *wavelet* analisadoras

Uma função *wavelet*  $\psi(t)$  deve satisfazer a condição de admissibilidade conforme Equação 3.1:

$$C_\psi = 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty, \quad (3.1)$$

Essa condição garante a inversibilidade da transformada, equivale, na maioria dos casos, a exigir que a integral de  $\psi(t)$  seja zero, garantindo também que a função possua forma oscilatória com média zero, apresentado na Equação 3.2:

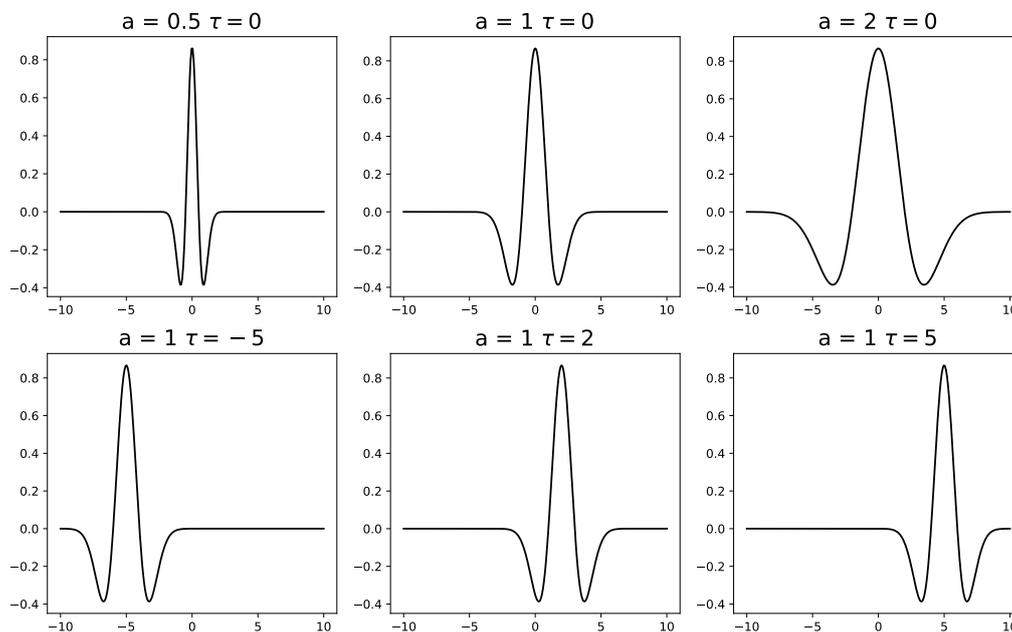
$$\int_{-\infty}^{\infty} \psi(t) dt = 0. \quad (3.2)$$

É necessário também a condição de energia unitária, que garante que a função *wavelet* possua localização espacial, por meio do suporte compacto, apresentada na Equação 3.3.

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

Na transformada *wavelet* as funções analisadoras são transladas com as variações do parâmetro de deslocamento temporal ( $\tau$ ) e dilatadas com variações de escalas ( $a$ ), para que seja possível realizar análises bem localizadas, respeitando o Princípio de Incerteza de Heisenberg (LAU; WENG, 1995; DOMINGUES et al., 2005). A Figura 3.1 apresenta um exemplo de variação dos parâmetros de uma função *wavelet* típica.

Figura 3.1 - Variação dos parâmetros  $\tau$  e  $a$  de uma função *wavelet*.



Fonte: Produção do Autor.

A escolha da função *wavelet* analisadora é um importante aspecto a ser considerado nas análises de dados. Existem inúmeras  $\psi$  possíveis, onde cada função analisadora destaca determinados aspectos do sinal. Para a análise de ciclos por meio da Fase, utilizam-se funções *wavelet* complexas, como a Paul e a Morlet. Para uma melhor localização temporal das variações, funções reais como a DOG e a chapéu mexicano são as mais indicadas (DOMINGUES et al., 2005; TORRENCE; COMPO, 1998).

### 3.1.1 Wavelet de Morlet

No contexto de geofísica, a *wavelet* analisadora mais utilizada é a função wavelet de Morlet. Sua principal característica, além de ser uma função complexa, é dada pelo formato de ondas planas dentro de um envelope Gaussiano (LAU; WENG, 1995). A *wavelet* de Morlet possui diversas Formulações, em 3.4 apresenta-se a Formulação simplificada introduzido por Grossmann and Morlet (1984) para estudos de sinais geofísicos:

$$\psi(t) = \pi^{-\frac{1}{4}} (e^{i\omega_0 t} - e^{\frac{t^2}{2}}). \quad (3.4)$$

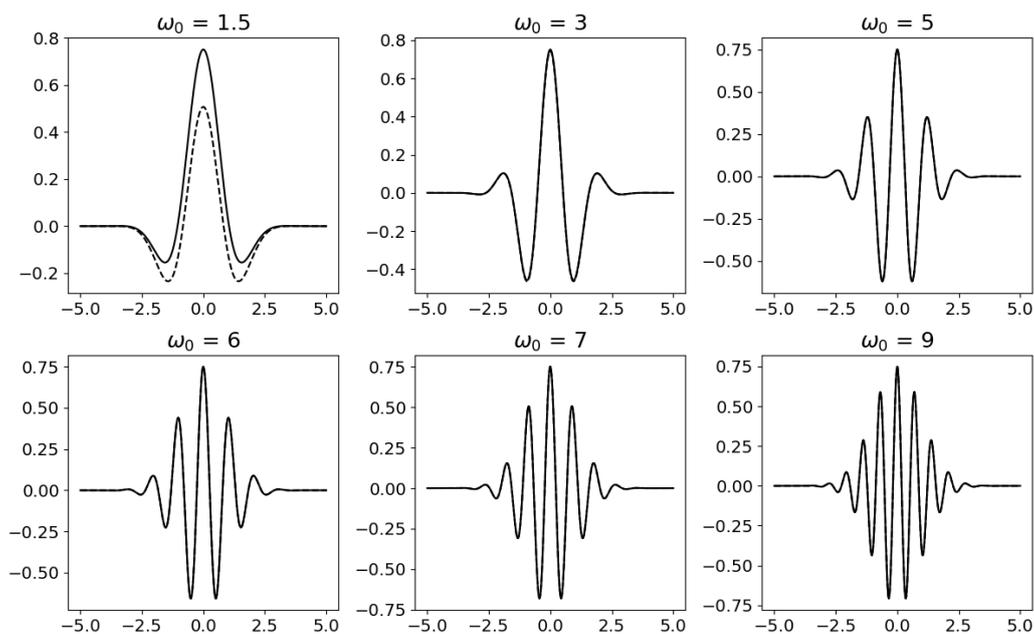
Essa formulação proporciona uma família de funções *wavelet* analisadoras dadas pela variação do parâmetro  $\omega_0$ . Entretanto, dependendo do valor dado a este parâmetro, perde-se a condição de admissibilidade, *i.e.* a função não atende mais o requisito

para ser considerada uma função *wavelet*. Normalmente adota-se  $\omega_0 = 6$ , dando a função de Morlet, equilíbrio entre localização no tempo e na frequência (FARGE, 1992; ANTOINE et al., 2008). Existe também uma formulação mais completa da *wavelet* de Morlet, dada por exemplo em 3.5. Esta fórmula ajusta melhor o decaimento da função de Morlet:

$$\psi(t) = \pi^{-\frac{1}{4}} (e^{i\omega_0 t} - e^{-\frac{\omega_0^2}{2}}) e^{-\frac{t^2}{2}}. \quad (3.5)$$

Na Figura 3.2 apresenta-se a família de funções de Morlet dada pela variação do parâmetro  $\omega_0$ , as funções com linha contínua são dadas pela Equação 3.4, as tracejadas são dadas pela Equação 3.5.

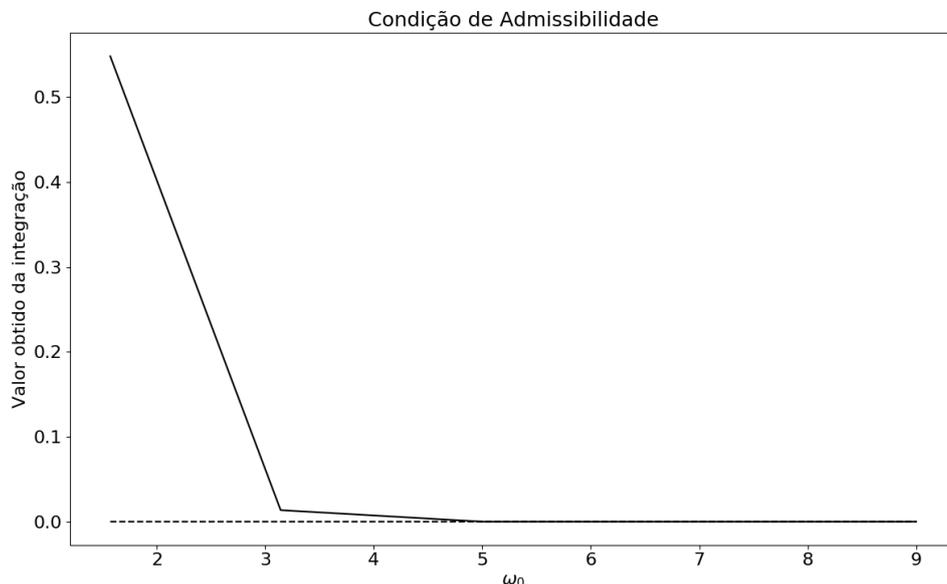
Figura 3.2 - Plotagem da variação do parâmetro  $\omega_0$  da parte real da *wavelet* de Morlet, utilizando a sua formulação simplificada (linha tracejada) e completa (linha contínua).



Fonte: Produção do Autor.

A Figura 3.3 apresenta a condição de admissibilidade em função da variação de  $\omega_0$  das funções *wavelets* de Morlet. Em 3.3 é possível observar que na formulação simplificada da *wavelet* de Morlet, conceitualmente considera-se efetivamente uma função *wavelet* analisadora apenas as funções cujo o parâmetro  $\omega_0$ , assumem valores maiores do que 5, enquanto na formulação completa, a condição de admissibilidade é satisfeita independente do valor utilizado para este parâmetro.

Figura 3.3 - Valor de integração da função *wavelet* de Morlet, simplificada (Equação 3.4) e completa (Equação 3.5). Valores igual a zero obedecem a condição de admissibilidade.



Fonte: Produção do Autor.

### 3.2 Transformada *wavelet* contínua

Um formalismo matemático hoje conhecido como a transformada *wavelet* contínua foi proposto por Morlet (1983), Grossmann and Morlet (1984) para análise de dados geofísicos, mostrava-se como uma forma de aperfeiçoar a Transformada de Fourier nos estudos de sinais não estacionários. Atualmente entende-se que os conceitos das *wavelets* eram desenvolvidos em diversas outras áreas devido a outras necessidades específicas, por exemplo, em processamento de sinais e na física do estado sólido, conforme descrito no livro de Antoine et al. (2008).

A CWT de uma função  $f(t)$  é dada pela formulação integral:

$$W_{\psi}^f(a, \tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \overline{\psi\left(\frac{t - \tau}{a}\right)} dt, \quad (3.6)$$

em que os parâmetros de escala ( $a$ ) e localização temporal ( $\tau$ ) assumem valores contínuos e  $W_{\psi}^f$  são denominados os coeficientes *wavelet*.

De forma similar a transformada discreta de Fourier, é possível calcular esses coeficientes dessa transformada discretamente, entretanto há uma alta complexidade

computacional envolvida nesse cálculo que é da ordem ( $\mathcal{O}(N^2)$ ), em que  $N$  é o número de elementos do vetor de dados discreto considerado (ANTOINE et al., 2008).

A normalização da transformada 3.6 é usualmente feita pela norma  $L^2$  (referente ao espaço das funções de quadrado integrável) como descrito Daubechies (1992). Dessa forma os coeficientes *wavelet* em módulo ao quadrado são indicadores da energia do sinal analisado. O termo energia é uma analogia a energia cinética usual da Física.

### 3.2.1 Ferramentas usuais da CWT

A análise de dados por meio da Transformada *Wavelet* Contínua é realizada nesta dissertação por meio de algumas das ferramentas apresentadas nesta seção.

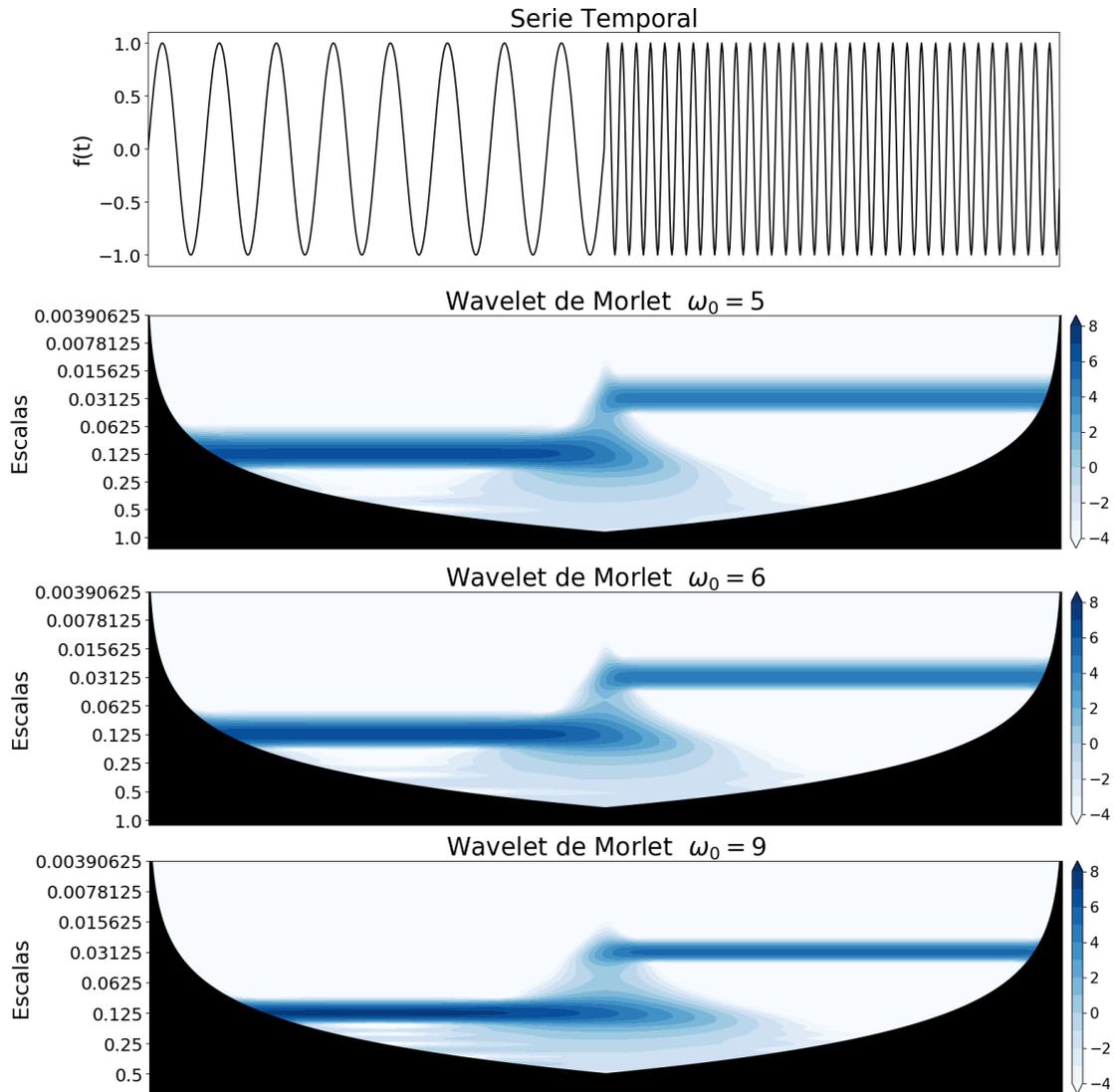
#### Escalograma

O escalograma é a visualização dos coeficientes *wavelets* obtidos diretamente da CWT. Em geral no escalograma é visualizado o módulo ao quadrado dos coeficientes *wavelet* i.e.  $|W_\psi^f(a, \tau)|^2$ . De forma prática, utiliza-se a visualização desses coeficientes por meio de imagens das matrizes de coeficientes obtidas, sendo os mapas de contorno ou padrões de cores as formas mais usuais de visualização. Um detalhe a ser destacado é que a escolha da paleta de cores pode alterar a facilidade de análise do escalograma. (FARGE, 1992). Pode-se ainda apresentar diferentes artifícios matemáticos para facilitar a análise realçando aspectos dos sinais analisados (ANTOINE et al., 2008). Neste trabalho apresenta-se as energias dos escalogramas em  $\log_2$ .

#### Exemplo de escalogramas utilizando as famílias da função *wavelet* de Morlet

Como exemplo, considera-se uma série temporal sintética dada por uma função seno que assume duas diferentes frequências ao longo do tempo. Na Figura 3.4 apresenta-se alguns escalogramas explorando as diferenças de análise ao se alterar as funções da família das *wavelets* de Morlet (formulação simplificada). Utiliza-se três valores distintos de  $\omega_0$  para a geração dos escalogramas. Observam-se nas três sub-figuras dos escalogramas, com esperado, estruturas com maior energia em duas escalas distintas que são alternadas no decorrer do tempo, representando o comportamento da função. As diferentes funções de Morlet proporcionam uma escolha no balanço da localização tempo-frequência apresentada no escalograma. Ao utilizar valores maiores em  $\omega_0$ , localiza-se de forma mais refinada as características frequenciais (em escala), em detrimento da localização temporal, essas características apresentam-se como uma escolha ao realizar as análises dos dados.

Figura 3.4 - Série temporal sintética e seus respectivos escalogramas relativos a diferentes famílias de Morlet.



Fonte: Produção do Autor.

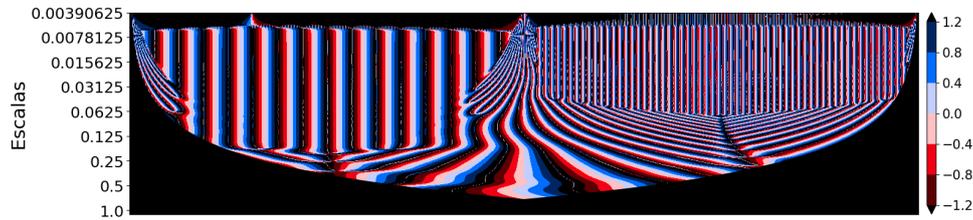
Com base nos coeficientes *wavelet*  $W_{\psi}^f(a, \tau)$  obtidos pela CWT com uma função analisadora complexa, é possível definir a fase  $\theta(a, \tau)$ , utilizando as partes reais e complexas desses coeficientes, apresentada na Equação 3.7.

$$\theta(a, \tau) = \arctan \left( \frac{\text{Im}(W_{\psi}^f(a, \tau))}{\text{Re}(W_{\psi}^f(a, \tau))} \right), \text{ em que } \theta \in (-\pi, \pi) \quad (3.7)$$

Para ilustrar a análise, a seguir apresenta-se a fase referente a série temporal  $f_1$  com

o parâmetro  $\omega_0 = 6$ .

Figura 3.5 - Fase dos coeficientes *wavelet* referente a função analisadora de Morlet (3.5 com  $\omega_0 = 6$ ).



Fonte: Produção do Autor.

Nesta figura são apresentados os valores de fase referentes aos coeficientes *wavelet* complexos.

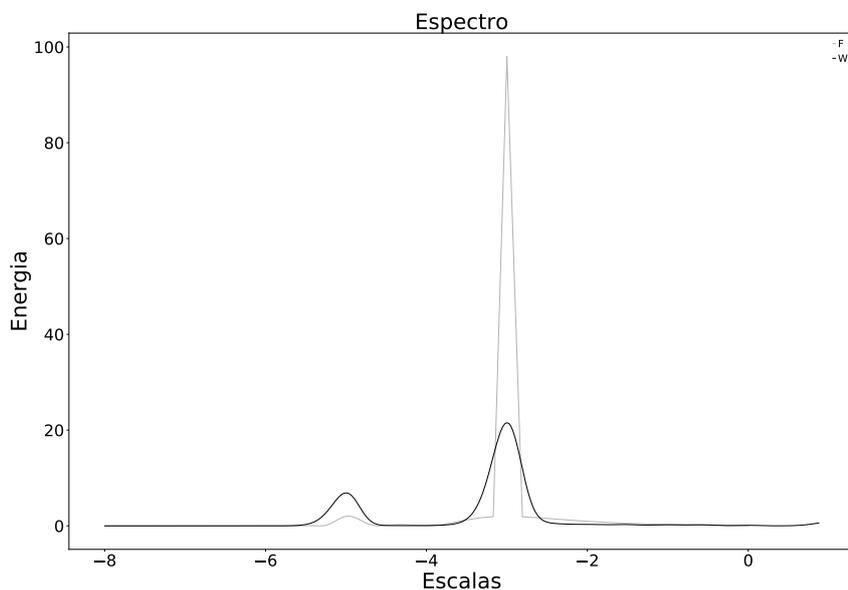
### Espectro *wavelet* global e aplicações

Interpreta-se o espectro *wavelet* global como a distribuição de energia do sinal em relação a escala, possui comportamento similar ao Espectro de Potências da análise Fourier, com características suavizadas e de menor amplitude. Esse espectro é dado pela integração dos coeficientes *wavelet* em relação ao parâmetro  $\tau$  (FARGE, 1992), assim:

$$\mathbb{S}_{\psi}^f(a) = \int_{-\infty}^{\infty} |W_{\psi}^f(a, \tau)|^2 d\tau \quad (3.8)$$

Para ilustrar o espectro global e compará-lo ao espectro de Fourier, utiliza-se como exemplo a CWT de  $f_1$ .

Figura 3.6 - Espectro *wavelet* global (em negrito) e espectro de Fourier.



Fonte: Produção do Autor.

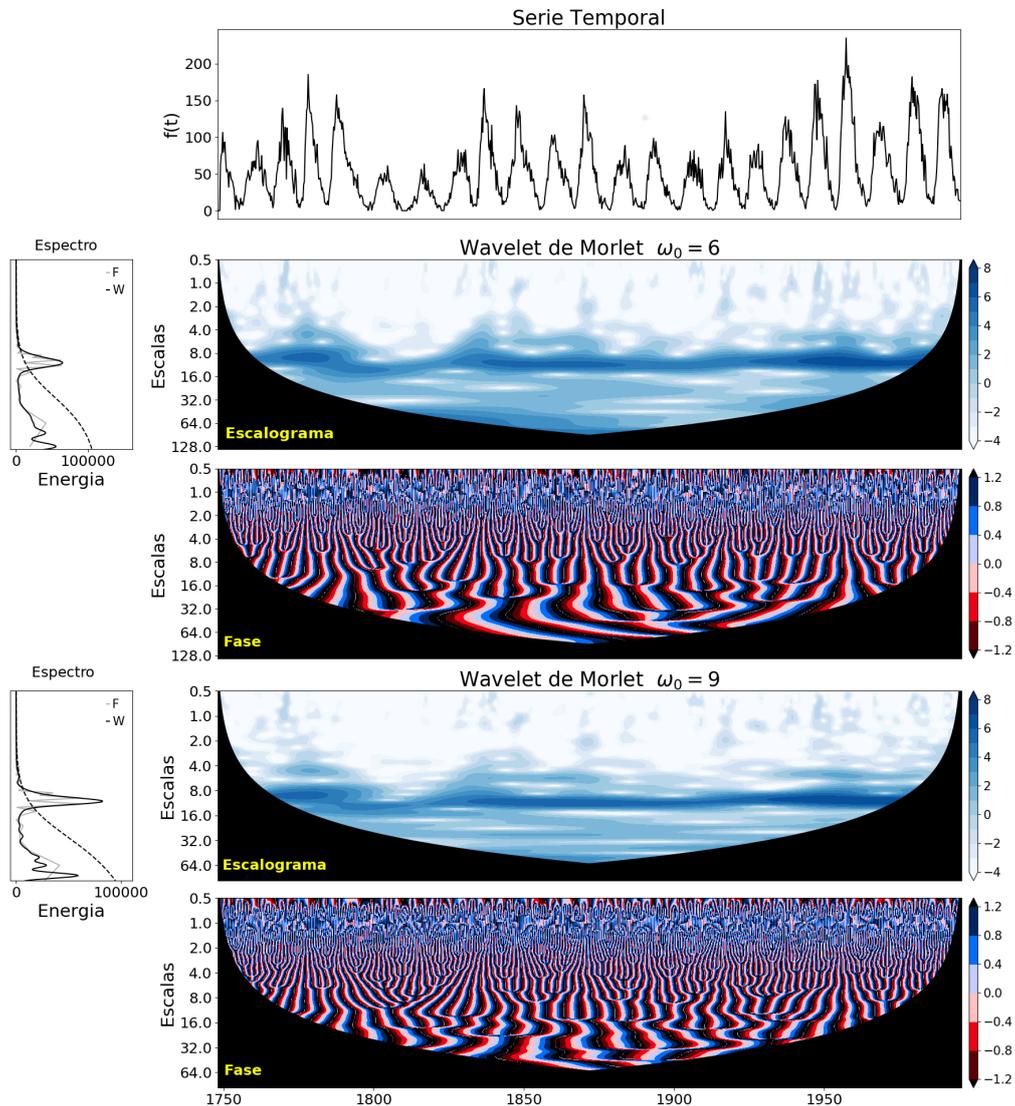
### 3.2.2 Análise e resultados em aplicações conhecidas

Nesta seção apresentam-se estudos preliminares das ferramentas da CWT utilizando séries temporais didáticas. Nesses exemplos utiliza-se a *wavelet* de Morlet com variações no parâmetro  $\omega_0$ .

#### Dados dos Ciclos de mancha solares

Esses dados contém observações dos anos 1748 à 1995 representando o número de manchas solares. O ciclo solar de 11 anos é o mais conhecido, e já documentado como um indicador de variações climáticas como por exemplo em Friis-Christensen and Lassen (1991) e geomagnéticas, com por exemplo em Russell and McPherron (1973). A Figura 3.7 apresenta as característica tempo-escala desta série temporal por meio das ferramentas da Transformada *Wavelet* Contínua.

Figura 3.7 - Série temporal de manchas solares, escalograma, fase e espectro.



Fonte: Produção do Autor.

O aspecto notável deste exemplo é a marcação com alta energia da escala que representa o conhecido ciclo de 11 anos, entre as escalas 8 e 16.

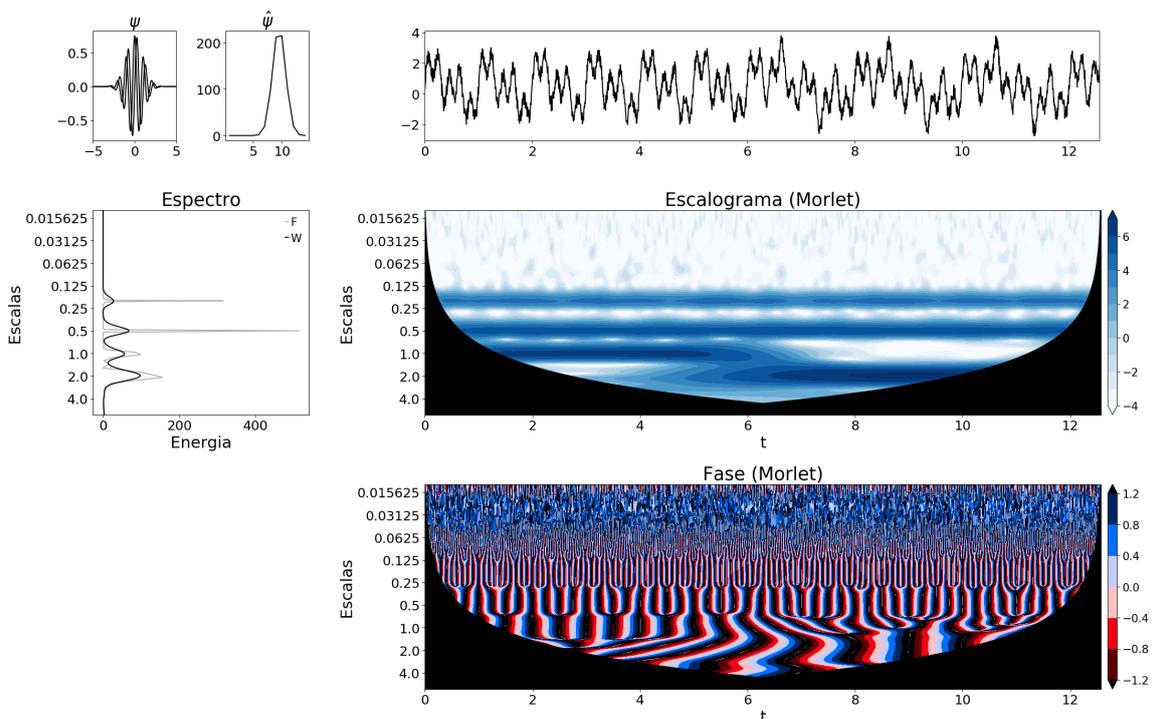
Ao variar o parâmetro  $\omega_0$  a localização do ciclo de 11 anos é refinada. Utilizando valores menores para  $\omega_0$  há um intervalo entre os anos de 1800 e 1840 em que os coeficientes possuem amplitude mais baixa na escala que representa esse ciclo, essa característica é menos destacada no escalograma com  $\omega_0 = 9$ . Pelo gráfico da fase é possível observar nesta escala que há sempre estruturas com comportamento diferenciado entre 1810 e 1840 em qualquer variação de  $\omega_0$ .

## Extração de bandas de frequências de um sinal não estacionário

Neste exemplo utiliza-se a série temporal  $f_2$ , dado em 3.9, para exemplificar o conceito de filtragem por escala com a CWT. Esse sinal é composto pela soma de três variações em diferentes frequências acrescidos de um ruído aleatório. A partir de um determinado momento, uma das variações se comporta de forma diferente, conforme apresentado na Figura 3.8.

$$f_2(t) = \begin{cases} \text{sen}(\pi t) + \text{sen}(4\pi t) + \text{sen}(10\pi t) + \epsilon & t < 2\pi \\ \text{sen}(2\pi t) + \text{sen}(4\pi t) + \text{sen}(10\pi t) + \epsilon & \text{caso contrário} \end{cases}, \quad (3.9)$$

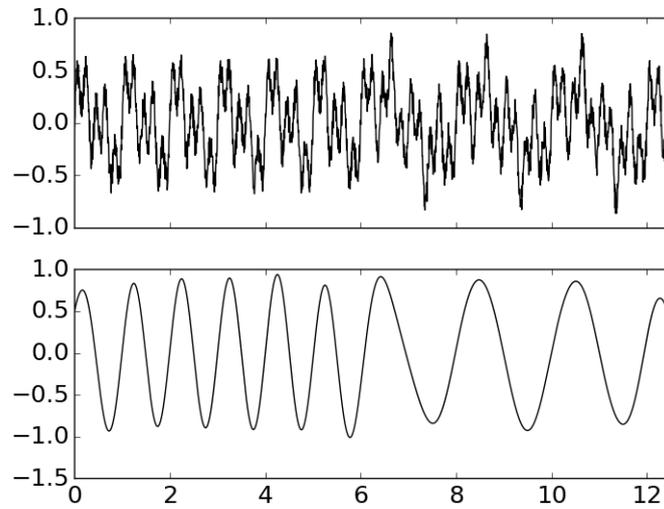
Figura 3.8 - Série temporal  $f_2$ , escalograma, espectros e fase.



Fonte: Produção do Autor.

Por meio do escalograma localiza-se essa componente que se diferencia ao longo do tempo entre as escalas 1 e 4. Para observar isoladamente esse comportamento, isola-se apenas essas escalas para produzir um sinal sintético. A Figura 3.9 apresenta o resultado dessa extração.

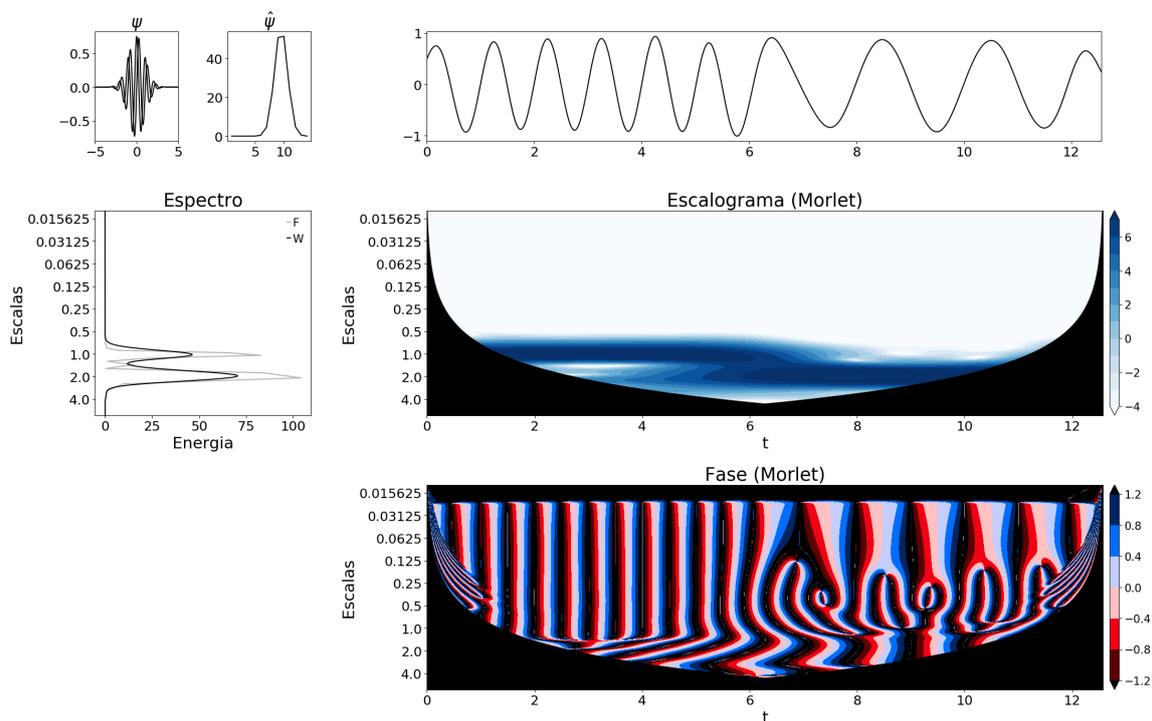
Figura 3.9 - Série temporal original e série temporal derivada.



Fonte: Produção do Autor.

Para verificar se a extração foi realizada corretamente, apresenta-se as características tempo-escala desse novo sinal na Figura 3.10, onde verifica-se que apenas as escalas de interesse se mantiveram.

Figura 3.10 - Série temporal derivada de  $f_2$ , escalograma, espectros e fase.



Fonte: Produção do Autor.

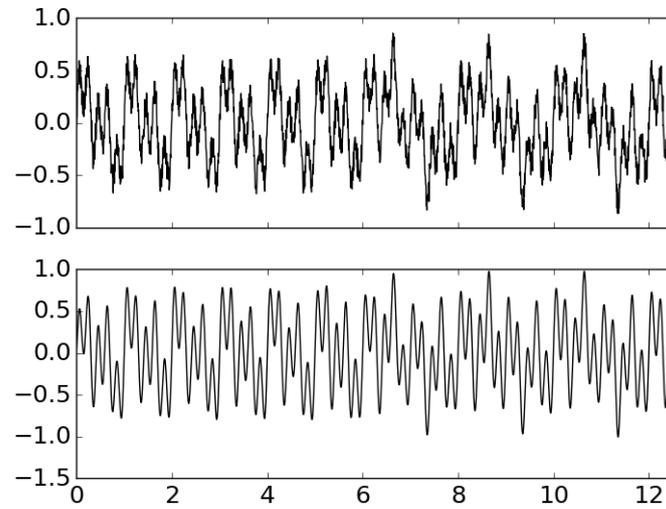
O processo de extração de sinal se dá por meio cálculo da CTW, após observar as características tempo-escala no escalograma, define-se o intervalo de extração, atribui-se o valor 0 para todas as outras escalas e inverte-se a CWT para obtenção de um sinal filtrado.

### Filtragem de ruídos de um sinal

A série temporal  $f_2$  possui um ruído sintético aleatório inserido propositalmente. Em alguns conjuntos de dados, esses ruídos podem ser provenientes de diferentes origens, podendo causar interferência e até enviesando análises. Neste exemplo utiliza-se a CWT para remoção desses ruídos sintéticos.

A Figura 3.8 apresenta o escalograma e os espectros associados ao sinal  $f_2$ , é possível observar que os ruídos estão presentes nas altas frequências deste sinal. A Figura 3.11 apresenta o resultado desta filtragem.

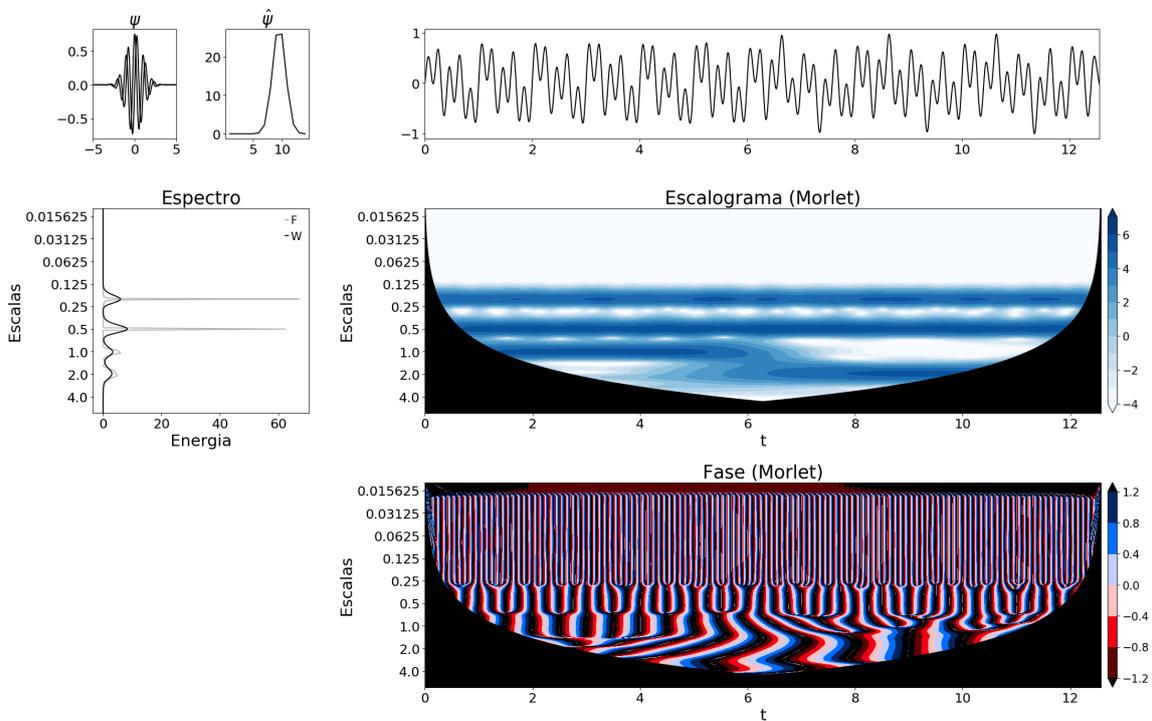
Figura 3.11 - Série Temporal  $f_2$  Original e Filtrada.



Fonte: Produção do Autor.

A Figura 3.12 apresenta as características tempo-escala da versão filtrada de  $f_2$ .

Figura 3.12 - Série temporal  $f_2$  filtrada, escalograma, espectros e fase.



Fonte: Produção do Autor.

### 3.2.3 Outras ferramentas da transformada *wavelet* contínua

Dentro do contexto de análise *wavelet* existem ferramentas que possibilitam análises de diversas características tempo-escala, como a entropia *wavelet* que apresenta a desordem de um sinal e também as ferramentas de estudos de relações em tempo-escala entre dois sinais, como a correlação por escala, a coerência *wavelet* e transformada *Wavelet* Cruzada. A seguir apresenta-se um resumo dessas ferramentas em algumas aplicações ilustrativas no contexto de ciências espaciais. Um estudo dos efeitos da suavização do escalograma para melhor representar a coerência *wavelet* é apresentado no Apêndice A.

#### 3.2.3.1 Entropia *wavelet*

Na física, o conceito de entropia é utilizado para medir a desordem de um sistema, em analogia à esse conceito, a entropia *wavelet* é utilizada para caracterizar a desordem de um sinal. A entropia *wavelet* é mínima quando o sinal apresenta atividade organizada; e máxima quando há grande números de processos influenciando um determinado tempo  $\tau$  de um sinal (SELLO, 2003; QUIROGA et al., 2001). A formulação da entropia *wavelet* é dada por:

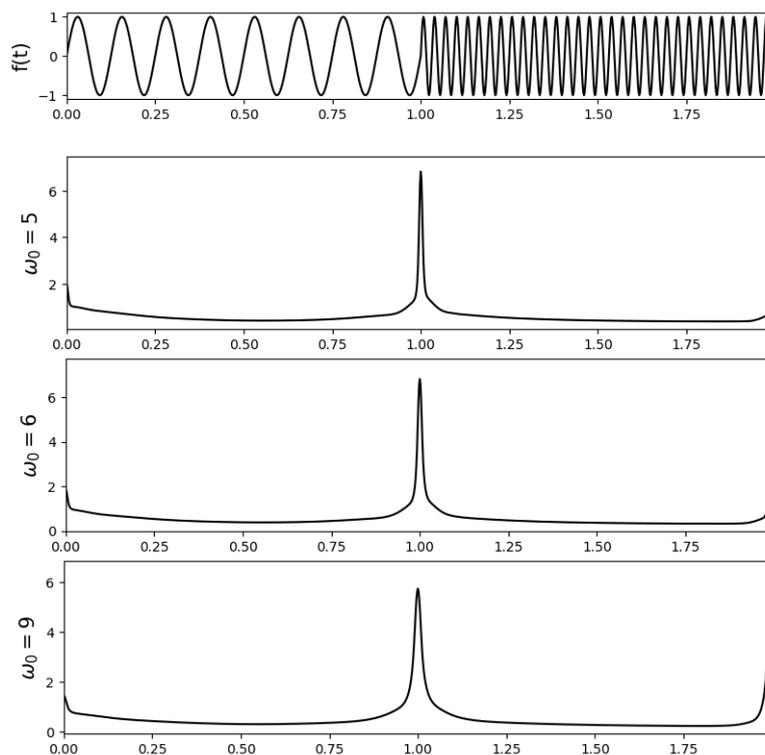
$$S_{\psi}(\tau) = \int_0^{\infty} -P_{\psi}(a, \tau) \ln(P_{\psi}(a, \tau)) da \quad (3.10)$$

em que,

$$P_{\psi}(a, \tau) = \frac{|W_{\psi}^f(a, \tau)|^2}{\int_{-\infty}^{\infty} |W_{\psi}^f(a, \tau)|^2 da} \quad (3.11)$$

Para apresentar a entropia *wavelet* utiliza-se a série temporal  $f_1$  com variação do parâmetro  $\omega_0$  da *wavelet* de Morlet.

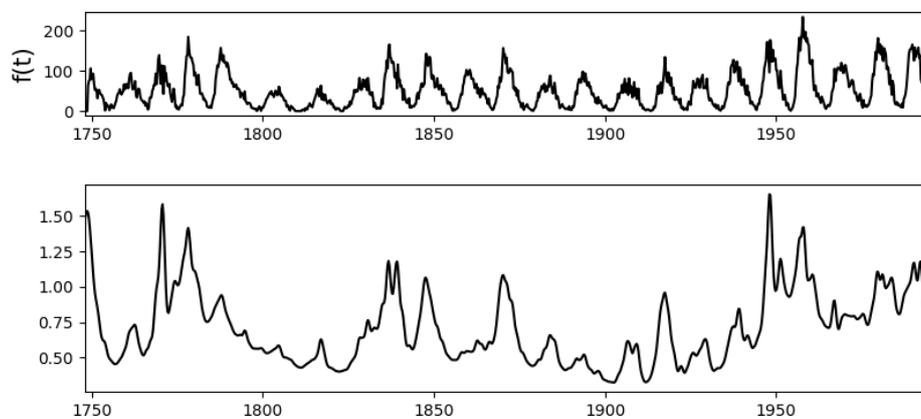
Figura 3.13 - Entropia *wavelet* da série temporal  $f_1$ .



Fonte: Produção do Autor.

Aplica-se o cálculo da entropia *wavelet* na série temporal de manchas solares (*Sunspot*) com a função *wavelet* de Morlet com  $\omega_0 = 6$ .

Figura 3.14 - Entropia *Wavelet* das série de manchas solares.



Fonte: Produção do Autor.

Observa-se na Figura 3.14 que as maiores medidas de entropia estão localizadas nas maiores amplitudes do sinal. Conclui-se que além da sobreposição de escalas, a amplitude do sinal também influencia a medida de entropia de um sinal.

### 3.2.3.2 Transformada *Wavelet* Cruzada - XWT

Ao analisar dados dentro de um determinado contexto, regularmente se encontra a necessidade de avaliar as relações entre dois ou mais diferentes sinais. Em análise *wavelet* existem ferramentas que possibilitam realizar esse tipo de análise dentro do domínio tempo-escala.

Uma forma de analisar duas séries temporais no contexto tempo-escala é a Transformada *Wavelet* Cruzada (XWT), em conjunto com o escalograma cruzado (HUDGINS et al., 1993). A formulação da XWT é dada por:

$$W_{\psi}^{f,g}(a, \tau) = W_{\psi}^f(a, \tau) \overline{W_{\psi}^g(a, \tau)}. \quad (3.12)$$

O escalograma cruzado  $|W_{\psi}^{f,g}(a, \tau)|^2$  apresenta a relação de dois sinais em uma matriz de variação de cores.

### 3.2.3.3 Coerência *Wavelet*

No contexto de processamento de sinais, coerência *wavelet* é uma medida de correlação em tempo-escala de dois sinais não-estacionários (LABAT, 2005). A formulação da coerência *wavelet* dada na Equação 3.13 foi apresentada inicialmente em Torrence and Webster (1999). Nessa formulação utiliza-se uma suavização dos coeficientes *wavelet* obtidos dos sinais  $f$  e  $g$ .

$$R^{f,g}(a, \tau) = \frac{|S(W_{\psi}^{f,g}(a, \tau))|}{[S(|W_{\psi}^f(a, \tau)|^2)S(|W_{\psi}^g(a, \tau)|^2)]^{\frac{1}{2}}} \quad (3.13)$$

Na prática, existem várias possibilidades para a função de suavização  $S$ , essas escolhas podem influenciar na localização tempo-escala dos fenômenos representados. Neste trabalho utiliza-se o filtro Gaussiano com  $\sigma = 6$ .

Calcula-se também a fase obtida da coerência *wavelet* conforme apresentado a seguir 3.14,

$$\phi^{f,g}(a, \tau) = \arctan \left( \frac{\text{Im}(R^{f,g}(a, \tau))}{\text{Re}(R^{f,g}(a, \tau))} \right). \quad (3.14)$$

Nesta ferramenta analisa-se se os sinais estão com comportamento em fase para

determinada escala, e quando não estão, apresenta uma relação de comportamento de um sinal em relação ao outro (AGUIAR; SOARES, 2014). Por exemplo:

- se  $\phi^{f,g}(a, \tau) = 0$  as séries se movem em conjunto na escala  $a$  no tempo  $\tau$ ;
- se  $\phi^{f,g}(a, \tau) \in (0, \frac{\pi}{2})$  as séries movem em fase na escala  $a$  no tempo  $\tau$ ;
- se  $\phi^{f,g}(a, \tau) \in (-\frac{\pi}{2}, 0)$   $f$  está adiantada  $g$  na escala  $a$  no tempo  $\tau$ ;
- se  $\phi^{f,g}(a, \tau) \in (\frac{\pi}{2}, \pi)$   $g$  adiantada em relação a  $f$  no tempo  $\tau$ ; na escala  $a$ ;
- a diferença de fase em  $\pi$  ou  $-\pi$  apresenta uma relação anti-fase na escala  $a$  no tempo  $\tau$ .

### 3.2.3.4 Correlação por escala

Outra forma de estudar as relações em diferentes escalas, entre dois sinais quaisquer  $f$  e  $g$  é a *Correlação por escala*  $\mathbb{C}_{S_\psi}(a)$ , dada na Equação 3.15.

$$\mathbb{C}_{S_\psi}(a) = \frac{\int_{-\infty}^{\infty} \mathbb{W}_\psi^g(a, \tau) \overline{\mathbb{W}_\psi^f(a, \tau)} d\tau}{[S_\psi^f(a) S_\psi^g(a)]^{\frac{1}{2}}}, \quad (3.15)$$

o termo  $\mathbb{W}_\psi^f(a, \tau)$  é expresso por:

$$\mathbb{W}_\psi^f(a, \tau) = |W_\psi^f(a, \tau) - \widetilde{W_\psi^f(a, \tau)}|, \quad (3.16)$$

$\widetilde{W_\psi^f(a, \tau)}$  é dada pela média de  $W_\psi^f(a, \tau)$  em  $\tau$ .

Dentro do domínio *wavelet* o valor de  $\mathbb{C}_{S_\psi}(a)$  apresenta em quais escalas há as maiores relações (FRICK et al., 2001).

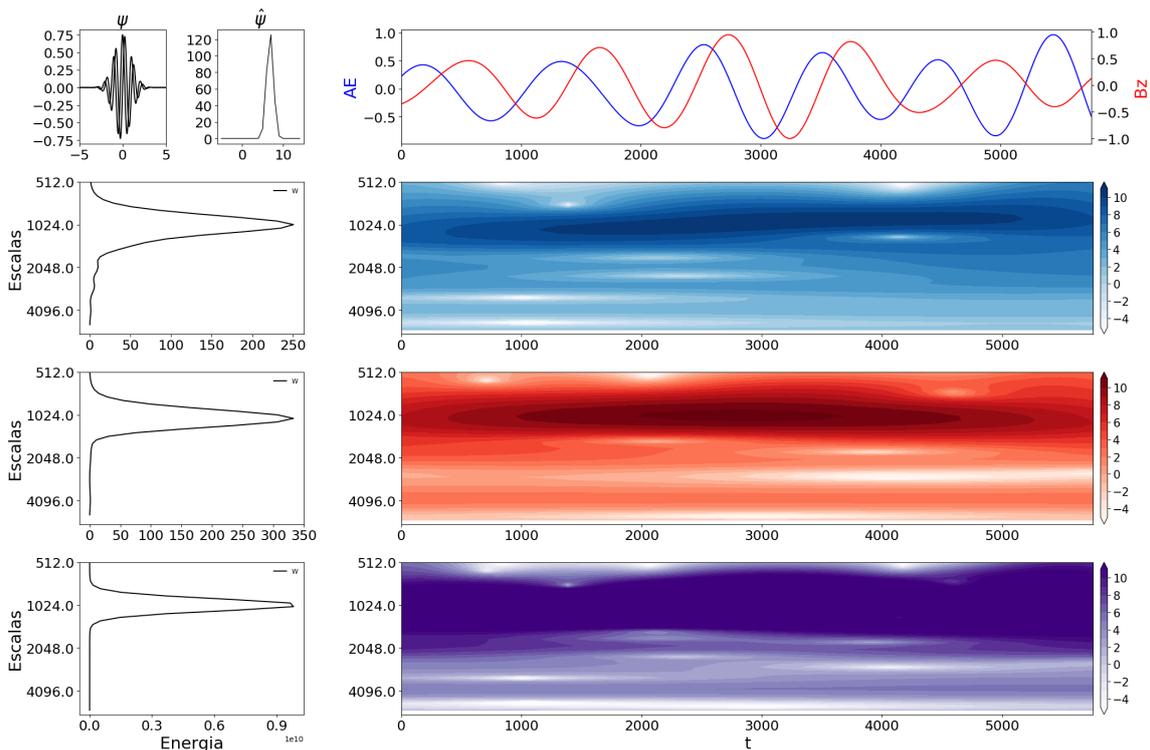
### 3.2.4 Análises preliminares das ferramentas de estudo de relações tempo-escala entre séries temporais

Nesta seção apresentam-se estudos de relações entre séries temporais sintéticas.

### 3.2.4.1 Resultados parciais: relações entre dados do meio interplanetário e de superfície

Em Seo et al. (2018), realizou-se um estudo de caso das relações tempo-escala de dados da componente  $B_z$  do campo magnético do meio interplanetário e do índice geomagnético AE em busca de possíveis evidências do acoplamento Sol-Terra em eventos HILDCAA utilizando ferramentas de análise de relações tempo-escala. Nesta aplicação, criam-se séries temporais simplificadas derivadas das escalas dentro do período do evento HILDCAA. A Figura 3.15 ilustra essa utilização, em que se lida com as séries temporais  $B_z$  e AE. Mostram-se os escalogramas de  $B_z$  e AE e também o escalograma cruzado.

Figura 3.15 - Séries temporais  $B_z$  e AE em que se evidencia a escala de 16 horas nos escalogramas de  $B_z$ , AE e Cruzado.

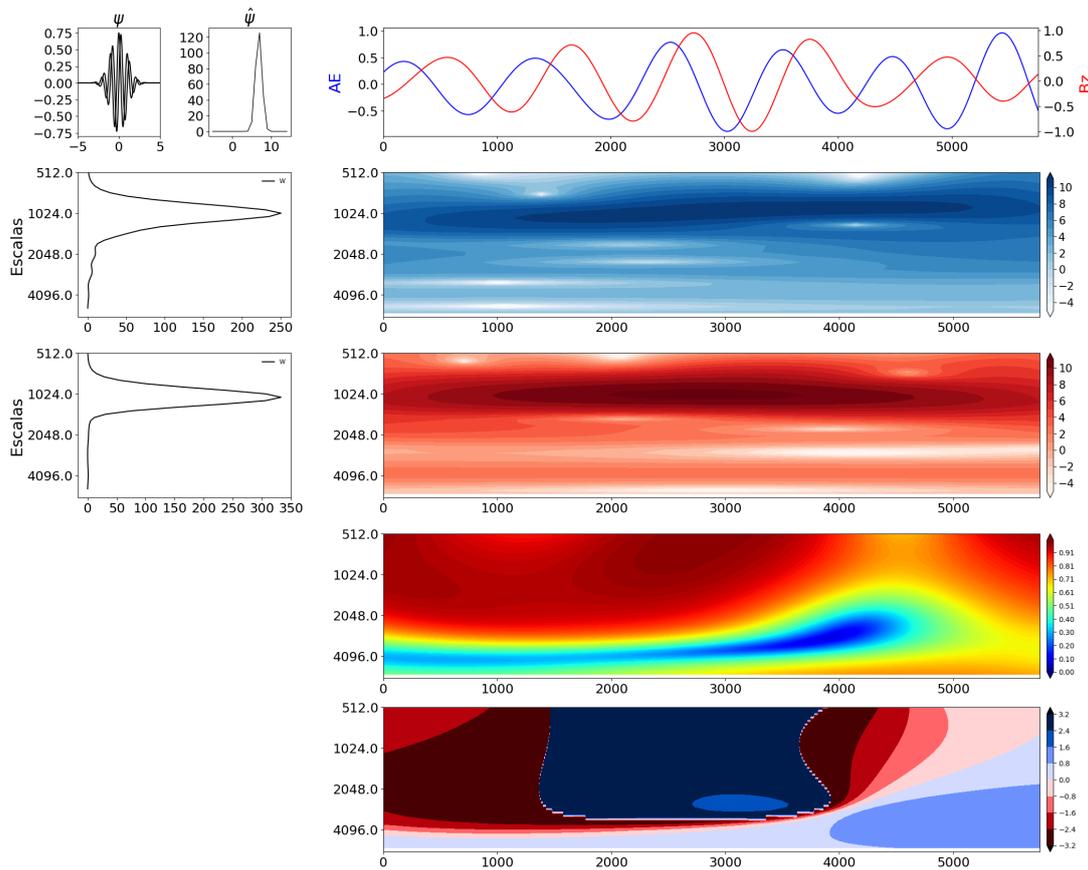


Fonte: Produção do Autor.

Observa-se por meio do escalograma cruzado que há uma estrutura entre as escalas 16 e 32 existente em ambos os sinais, ou seja, durante quase todo o domínio do tempo há uma relação tempo-escala acontecendo nesta escala.

A Figura 3.16 apresenta os escalalogramas, de cima para baixo, do  $AE$  e do  $B_z$  e da coerência *wavelet* e a correlação por escala entre  $B_z$  e  $AE$ .

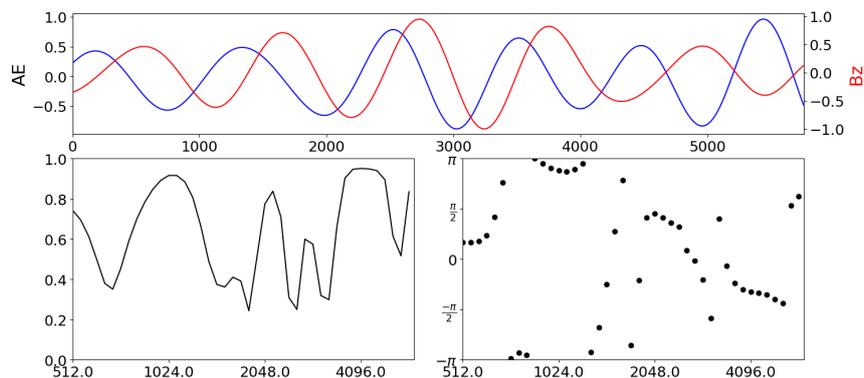
Figura 3.16 - Séries temporais  $B_z$  e  $AE$ , Escalograma de  $B_z$ , Escalograma de  $AE$ , Coerência *wavelet* e diferença de fase .



Fonte: Produção do Autor.

A coerência também destaca uma grande estrutura de relação tempo-escala, as principais estruturas localizadas próximas a escala 1024. Observa-se na diferença de fase uma grande mudança da condução da coerência entre os momentos 1200 e 4000. Analisa-se a correlação por escala e fase na Figura 3.17.

Figura 3.17 - Correlação por escalas de  $B_z$  e AE módulo e fase.



Fonte: Produção do Autor.

Observa-se que a correlação por escalas apresenta picos próximos as escalas 1024, 2048 e 4096. Os valores de correlação das escalas próximas a 2048 e 4096 podem ser desconsideradas pois as componentes nestas foram descartadas no processo de filtragem.

Observa-se na fase da correlação que os sinais estão com fortes mudanças de fase, nas regiões próximas as principais estruturas de ambos os sinais, a correlação passa subitamente de uma relação de quadratura para uma relação em fase.

### 3.2.5 Implementação da CWT e ferramentas

No contexto da CWT, neste trabalho utilizou-se a implementação do pacote **kPyWavelet** (disponível em: <https://github.com/wafels/kPyWavelet>) como base para o cálculo da transformada. Os cálculos das funções de relação tempo-escala e entropia *wavelet* foram desenvolvidos com base no trabalho de Domingues et al. (2016).

## 4 ANÁLISE POR TRANSFORMADAS *WAVELET* DISCRETAS

Neste capítulo, como a segunda parte para estabelecer a metodologia integrada de análise de sinais, introduz-se e estuda a Transformada *Wavelet* Discreta (DWT, Discrete *Wavelet* Transform), considerando suas características. A DWT é utilizada na decomposição ortogonal ou não dos dados para obtenção de séries temporais com contribuições reduzidas em escalas específicas, sendo assim uma candidata natural a obtenção de séries sintéticas representativas em certas escalas ou de um conjunto delas. Estabelecem-se e demonstram-se, igualmente, para a técnica algumas de suas potenciais aplicações.

A transformada *wavelet* discreta é caracterizada pela adoção de parâmetros de escala e translação discretos, usualmente considera-se,  $a = 2^j$  e  $b = 2^j k$ , em que  $j$  é denominado nível,  $k$  posição e os coeficientes *wavelets* são denotados por  $d_k^j$ . Essa notação para os coeficientes *wavelets*, como mais adiante será detalhado, se dá por eles representarem detalhes de mais alta frequência ou erros de aproximação do sinal em uma representação multi-nível pois representam a diferença do sinal em dois níveis consecutivos de aproximação. Formalmente, a transformada *wavelet* discreta é dada pela transformação integral

$$d_k^j = 2^j \int_{-\infty}^{\infty} f(t) \psi(2^j t - k) dt. \quad (4.1)$$

Essa transformada discreta pode ter várias características específicas que auxiliam o estudo de sinais não-estacionários.

Do ponto de vista prático a construção de funções *wavelets* que sejam de interesse a análises específicas e de transformadas discretas que possuam algoritmos rápidos é dada por uma ferramenta matemática conhecida como análise multirresolução.

A seguir apresenta-se a definição de uma análise multirresolução e das características para a obtenção dos coeficientes *wavelets* em transformadas discretas sob esse paradigma conhecidas como decimadas e sem decimação, utilizadas nesse trabalho.

### 4.1 Definição de uma Análise Multirresolução (MR)

Funções  $f \in L^2(\mathcal{R})$  também podem ser aproximadas pelas suas projeções em  $V^j$  e  $W^j$ , respectivamente,

$$\mathcal{P}^j f(t) = \sum_k c_k^j \phi_k^j(t), \quad \mathcal{Q}^j f(t) = \sum_k d_k^j \psi_k^j(t). \quad (4.2)$$

em que os coeficientes escala como  $c_k^j$  são definidos pelo produto interno do  $L^2$  entre a função em estudo e a função escala e *wavelet*, respectivamente,

$$c_k^j = \langle f, \phi_k^j \rangle, \quad d_k^j = \langle f, \psi_k^j \rangle. \quad (4.3)$$

Dessa forma tem-se que os coeficientes *wavelets* contém a diferença de informação entre os níveis  $j$  e  $j+1$  e por isso representam o erro de aproximação entre esse níveis ou detalhes de mais alta frequência encontrados apenas no nível mais refinado.

Conseqüentemente, em multinível para  $j_0 > j+1$ , têm-se então a decomposição

$$V^{j+1} = V^{j_0} + W^0 + \dots + W^{j+1}, \quad (4.4)$$

dessa forma,

$$\sum_k c_k^j \phi_k^j(t) = \sum_k c_k^{j+1} \phi_k^{j+1}(t) + \sum_k d_k^{j+1} \psi_k^{j+1}(t) \quad (4.5)$$

em multinível passa-se a ter a seguinte formulação

$$\sum_k c_k^j \phi_k^j(t) = \sum_k c_k^{j_0} \phi_k^{j_0}(t) + \sum_{m=j_0}^{j+1} d_k^m \psi_k^m(t) \quad (4.6)$$

Para efetuar essas operações, têm-se a Transformada *Wavelet* Discreta (DWT) e sua inversa (iDWT), ou seja,

$$c^{j+1} \rightleftharpoons c^{j_0}, d^{j_0}, \dots, d^{j+1}. \quad (4.7)$$

## 4.2 Algoritmos

Algoritmos de decomposição, relacionados a DWT, e reconstrução, relacionados a DWT inversa, são conhecidos atualmente como *algoritmos de Mallat*, referente ao trabalho seminal desenvolvido por Mallat (MALLAT, 1989). Atualmente pode-se também identificar esses algoritmos como transformada *wavelet* decimada, já que o processo de decimação é empregado, como descrito a seguir no texto. Mais tarde, variantes da DWT decimada surgiram, uma série delas são as transformadas discretas não decimadas. Entre uma das inicialmente propostas nesse contexto e uma das mais famosas na área de sinais está a *Maximum Overlap Discrete Wavelet Transform* (MODWT), discutida na sequência deste capítulo .

## Transformada decimada

No processo de decomposição da transformada *wavelet* discreta, os coeficientes  $c_k^{j+1}$  e  $d_k^{j+1}$  são obtidos dos coeficientes  $c_k^j$  por meio de algumas manipulações das relações de escala de  $\phi$  e *wavelet*  $\psi$  de tal forma que

$$c_k^{j+1} = \sum_{\ell=0}^{L-1} h(\ell - 2k) c_\ell^j \quad (4.8)$$

e

$$d_k^{j+1} = \sum_{\ell=0}^{L-1} g(\ell - 2k) c_\ell^j, \quad (4.9)$$

em que  $L$  é um número par que representa o número dos coeficientes não-zero do filtro  $h$ , sendo que  $h(0)$  e  $h(L-1)$  são diferentes de zero e  $h(\ell) = 0$  para  $\ell < 0$  e  $\ell \geq L$ . As equações apresentadas possuem ainda uma normalização que do ponto de vista computacional é incluída nos filtros utilizados.

Um exemplo de filtros escala  $h$  utilizados nesse trabalho para o caso da DWT decimada são os filtros ortogonais da família de Daubechies de ordem 8, conhecidos com *Daub 8* apresentados na Tabela 4.1. Esses filtros geram *wavelet* de suporte compacto com fase extrema e o máximo número de momentos nulos compatível com seu suporte. Neste caso apresentado a normalização do filtro é  $\sqrt{2}$ .

As Equações 4.8 e 4.9 compõe o algoritmo piramidal para a transformada discreta *wavelet* ortogonal <sup>1</sup>, conhecido como algoritmo de Mallat, esquematizado na Figura 4.1. Nessa esquema, o quadrado indica convolução com um filtro. O nome do filtro está indicado no interior do quadrado. O símbolo  $2 \downarrow$  indica um *downsampling* ou decimação, i.e., são eliminados os elementos de índice ímpar. O símbolo  $2 \uparrow$  indica um *upsampling*, i.e., são acrescentados componentes com valor zero entre as componentes do vetor original. Assim, dado um vetor  $c^j = (c_k^j)$ , obtém-se um novo vetor  $c^j$  pela aplicação de uma convolução entre  $c^j$  e o filtro  $h$  seguida de decimação. A diferença de informação entre  $c^j$  e  $c^{j+1}$ , ou seja, os detalhes de alta frequência, ficam contidos no vetor  $d^{j+1}$  obtido pela convolução entre  $c^{j+1}$  e o filtro  $g$  seguida de uma decimação. Os  $d_k^j$  correspondem aos coeficientes das expansões nas bases *wavelet*. Esse processo pode ser repetido outras vezes para obter os coeficientes  $d_k^{j+1}$ ,  $d_k^{j+2}$ ,  $d_k^{j+3}$ ,  $\dots$ . O processo de recuperação do sinal, a partir de sua análise MR, é semelhante ao descrito anteriormente. Neste caso, o cálculo dos coeficientes  $c_k^j$  em

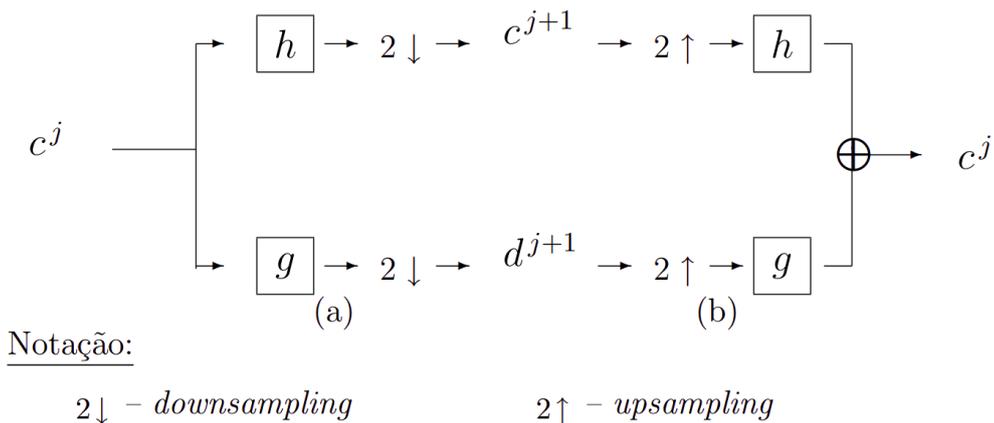
<sup>1</sup>Para o caso não-ortogonal tem-se a soma direta (+) e neste caso ortogonal tem-se a soma ortogonal ( $\oplus$ ).

termos de  $c_k^{j+1}$ ,  $d_k^{j+1}$  é dado por

$$c_{k,l}^j = \sum_{k^l} h(k - 2k^l) c_{k^l}^{j+1} + \sum_{k^l} g(k - 2k^l) d_{k^l}^{j+1} \quad (4.10)$$

Esse algoritmo também pode ser representado pela estrutura de bancos de filtros da parte direita dessa figura. Primeiro aplica-se um *upsampling* nos vetores  $c^{j+1}$  e  $d^{j+1}$ , seguido da convolução com o filtro  $h$  e  $g$ , respectivamente. A seguir, somam-se os resultados obtidos e encontra-se o vetor  $c^j$ . Ao final do processo de decomposição, o armazenamento dos coeficientes *wavelet* e da escala é feito no mesmo vetor de dados inicial e, dessa forma, não há necessidade de área física extra de armazenamento.

Figura 4.1 - Esquema da DWT decimada e ortogonal. Algoritmo de Decomposição (a) e de Reconstrução (b).



Fonte: Adaptado de Domingues (2001)

Uma descrição de uma possível implementação desses cálculos é apresentada no Algoritmo 1. Nele descreve-se a decomposição e reconstrução em um nível da DWT decimada ortogonal descrita no esquema da Figura 4.1. Se aplicarmos sucessivamente a decomposição tem-se a decomposição multi-nível dos vetores da DWT  $d^1, \dots, d^J$  e  $c^J$ . Na reconstrução calcula-se o vetor de dados original  $c^{j=0}$  utilizando os subvetores  $d^1, \dots, d^J$  e  $c^J$ . Esse algoritmo é adaptado do apresentado em (PERCIVAL; WALDEN, 2000), vale a pena destacar que lá a notação de encaixe dos espaços, da circulação dos filtros e da denominação dos filtros escala e *wavelet* são diferentes.

---

**Algoritmo 1:** Transformada decimada em um nível (periódica).

---

```
1: procedure DECOMPOSIÇÃO
    entrada: Vetor de dados  $c^j$ , Vetor do filtro escala  $h$ 
    saída : Vetores  $c^{j+1}$  e  $d^{j+1}$ 
     $N \leftarrow \#$  elementos do vetor  $c^j$ ;  $L \leftarrow \#$  elementos do filtro  $h$ 
    for  $\ell \leftarrow 0$  to  $L-1$  by 1 do
         $g_k \leftarrow (-1)^k h(-k+1)$  // pass-band filter
    end
     $J \leftarrow \log_2(N)$ 
     $M \leftarrow \frac{N}{2^J}$ 
    for  $k \leftarrow 0$  to  $M/2 - 1$  by 1 do
         $m \leftarrow 2k + 1$ 
         $d_k^{j+1} \leftarrow g_0 c_m^j$ 
         $c_k^{j+1} \leftarrow h_0 c_m^j$ 
        for  $\ell = 1, \dots, L-1$  do
             $m = m - 1$ 
            if  $m < 0$  then
                 $m \leftarrow M - 1$ 
            end
             $c_k^{j+1} \leftarrow c_k^{j+1} + h_\ell c_m^j$ 
             $d_k^{j+1} \leftarrow d_k^{j+1} + g_\ell c_m^j$ 
        end
    end

2: procedure RECONSTRUÇÃO
    entrada: Vetores  $c^{j+1}$ ,  $d^{j+1}$ , Vetor do filtro escala  $h$ 
    saída : Vetor de dados  $c^j$ 
     $M \leftarrow \#$  elementos do vetor  $c^{j+1}$  ou  $d^{j+1}$ ;
     $L \leftarrow \#$  elementos do filtro  $h$ 
     $\ell = -2$ ;  $m = -1$ 
    for  $k = 0, \dots, M - 1$  do
         $\ell \leftarrow \ell + 2$ ;  $m \leftarrow m + 2$ ;  $u \leftarrow k$ ;  $q \leftarrow 1$ ;  $p \leftarrow 0$ 
         $c_\ell^j \leftarrow g_q d_u^{j+1} + h_q c_u^{j+1}$ 
         $c_m^j \leftarrow g_p d_u^{j+1} + h_p c_u^{j+1}$ 
        if  $L > 2$  then
            for  $n = 1, \dots, L/2 - 1$  do
                 $u \leftarrow u + 1$ 
                if  $u \geq M$  then
                     $u \leftarrow 0$ 
                end
                 $q \leftarrow q + 2$ ;  $p \leftarrow p + 2$ 
                 $c_\ell^j = c_\ell^j + g_q d_u^{j+1} + h_q c_u^{j+1}$ 
                 $c_m^j = c_m^j + g_k d_u^{j+1} + h_k c_u^{j+1}$ 
            end
        end
    end
end
```

1 Fonte: Adaptada de (PERCIVAL; WALDEN, 2000, Sec. 4.6).

---

## Transformada não-decimada

A MODWT é uma variação da DWT não decimada proposta em meados dos anos 90, uma descrição mais detalhada desta transformadas é apresentada em (PERCIVAL; WALDEN, 2000, Cap.6). Diferentemente do processo da DWT, descrito nas Equa-

ções 4.8 e 4.9 e esquematizado na Figura 4.1, a MODWT não passa pelo processo de decimação (*downsampling*) e utiliza filtros específicos, como por exemplo os filtros Daubechies com a menor assimetria possível (do inglês, *Least Asymmetric filter*, LA) para essa decomposição, como os apresentados na Tabela 4.1. Nesse caso, por construção, eles possuem o maior número de momentos nulos.

De uma forma geral, numa formulação similar a apresentada nas Equações 4.8 e 4.9, os coeficientes escala e *wavelet* no algoritmo da MODWT podem ser descritos em dois níveis como:

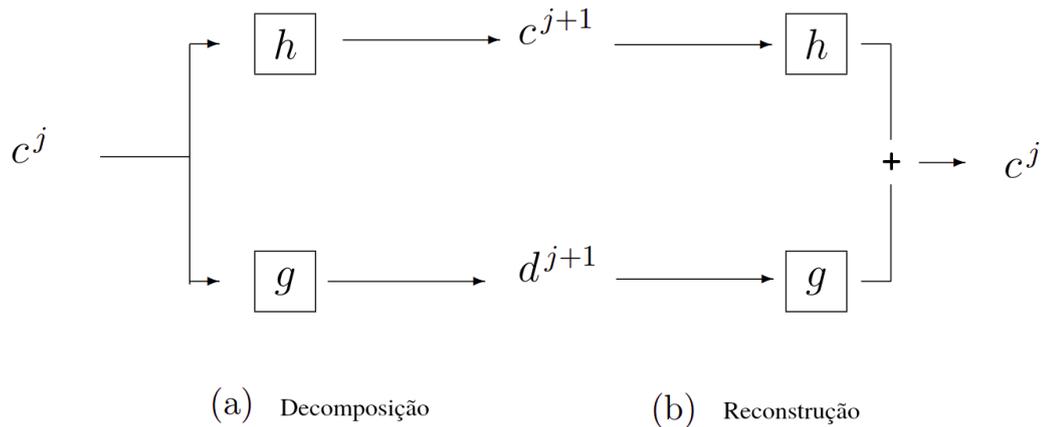
$$c_k^{j+1} = \sum_{\ell=0}^{L-1} h(\ell) c_{k-\ell}^j \quad (4.11)$$

e

$$d_k^{j+1} = \sum_{\ell=0}^{L-1} g(\ell) c_{k-\ell}^j. \quad (4.12)$$

De forma similar a Figura 4.1, representa-se o algoritmo de Mallat da MODWT na Figura 4.2 para dois níveis. Nesse caso a transformada não é ortogonal e não possui a decimação.

Figura 4.2 - Esquema da MODWT adaptado da DWT em Domingues (2001), baseada no algoritmo de Percival and Walden (2000). Algoritmo de decomposição (a) e de reconstrução (b).



Fonte: Produção do Autor.

Analogamente a DWT decimada, a decomposição não-decimada com a MODWT

é obtida inicialmente de  $\tilde{c}_k^{j=0}$ , para  $k = 0, \dots, N - 1$ , que representa a série de dados originais e do filtro escala para a MODWT representado pelo vetor  $\{h_\ell, \ell = 0, \dots, L - 1\}$ . Os coeficientes como  $d^{j=1}$  e  $c^{j=1}$  são calculados na decomposição. A reconstrução é o processo inverso, em que com os coeficientes  $d^1$  e  $c^1$  dão origem ao coeficientes  $c^0$ . O Algoritmo 2 apresenta a decomposição e reconstrução entre dois níveis:  $j \leftrightarrow j + 1$ . Neste caso o número de elementos dos vetores decompostos é o mesmo do vetor original. Para a decomposição em multi-nível utiliza-se o Algoritmo 2 recorrentemente para os níveis  $j = 1, j = 2, \dots, J$  obtendo-se os vetores de dados correspondentes aos coeficientes escala  $\tilde{c}_1, \dots, \tilde{c}_J$  e *wavelet*  $\tilde{d}_1, \dots, \tilde{d}_J$  da MODWT, em que  $J$  é último nível da decomposição, sendo  $N \geq 2^J$ . Na reconstrução calcula-se o vetor  $c^j$  utilizando os subvetores  $c^{\tilde{j}+1}$  e  $d^{\tilde{j}+1}$ .

---

**Algoritmo 2:** Transformada não decimada em um nível (periódica).

---

```

1: procedure DECOMPOSIÇÃO
    entrada: Vetor de dados  $c^j$ , Vetor do filtro escala  $h$ 
    saída : Vetores  $c^{j+1}$  e  $d^{j+1}$ 
     $N \leftarrow \#$  elementos do vetor  $c^j$ ;  $L \leftarrow \#$  elementos do filtro  $h$ 
    for  $\ell \leftarrow 0$  to  $L-1$  by 1 do
         $g_\ell \leftarrow (-1)^\ell h(L-1-\ell)$  // pass-band filter
    end
     $J \leftarrow \log_2(N)$ 
     $M \leftarrow N - 1$ 
    for  $k \leftarrow 0$  to  $M$  by 1 do
         $m \leftarrow k$ 
         $c_k^{j+1} \leftarrow h_0 c_m^j$ 
         $d_k^{j+1} \leftarrow g_0 c_m^j$ 
        for  $\ell = 1, \dots, L-1$  do
             $m = m - 2^j$ 
            if  $m < 0$  then
                 $m \leftarrow m \bmod N$ 
            end
             $c_k^{j+1} \leftarrow c_m^{j+1} + h_\ell c_m^j$ 
             $d_k^{j+1} \leftarrow d_m^{j+1} + g_\ell c_m^j$ 
        end
    end

2: procedure RECONSTRUÇÃO
    entrada: Vetores  $c^{j+1}$ ,  $d^{j+1}$ , Vetor do filtro escala  $h$ 
    saída : Vetor de dados  $c^j$ 
     $M \leftarrow N - 1$ ;
     $L \leftarrow \#$  elementos do filtro  $h$ 
    for  $k \leftarrow 0$  to  $M$  by 1 do
         $m \leftarrow k$ 
         $c_k^j \leftarrow h_0 c_k^{j+1} + g_0 d_k^{j+1}$ 
        for  $\ell = 1, \dots, L-1$  do
             $k \leftarrow k + 2^j$ 
            if  $m \geq N$  then
                 $m \leftarrow m \bmod N$ 
            end
             $c_\ell^j = c_\ell^j + g_\ell d_m^{j+1} + h_\ell c_m^{j+1}$ 
        end
    end

```

<sup>1</sup> Fonte: Adaptada de (PERCIVAL; WALDEN, 2000).

---

### 4.2.1 Comentários sobre os algoritmos e códigos utilizados

Em resumo as principais diferenças entre os esquemas apresentados é a não existência da operação de *downsampling* dos coeficientes e os filtros escolhidos. Conseqüentemente na MODWT os coeficientes gerados tem, então, o mesmo número de dados da série temporal original, mantendo determinados detalhes que poderiam ser perdidos no processo de decimação sob certas situações (NASON, 2008).

O algoritmo da MODWT tem a complexidade na ordem de  $\mathcal{O}(N \log_2 N)$ , tornando-

a mais custosa do que a DWT que é da ordem de  $\mathcal{O}(N)$ , mas ainda menos custosa do que a CWT da ordem de  $\mathcal{O}(N^2)$ , podendo ser considerada um meio termo no balanço custo *versus* redundância de informações (JACH et al., 2006).

Nos cálculos da DWT utiliza-se a linguagem python com as funções do pacote pywt, disponíveis na página da Internet <https://pywavelets.readthedocs.io/>. Adicionalmente, fez-se a implementação de uma camada de configuração de entrada de dados e da visualização dos coeficientes *wavelets* nas diversas escalas da decomposição. Para a utilização do código desenvolvido, criou-se uma estrutura de dados colunar no formato csv, contendo uma coluna para o tempo e o outro para os valores das quantidades físicas analisadas. As configurações de diretório, níveis de decomposição e filtros são realizadas no arquivo `config_file.json`.

Nos cálculos da MODWT utilizou-se também a linguagem python com funções do pacote modwtpy, disponíveis na página da Internet <https://github.com/pistononly/modwtpy>. Fez-se a implementação da camada de configuração de entrada de dados e da visualização dos coeficientes *wavelets* nos diversos níveis de decomposição. Utiliza-se nesta implementação a mesma estrutura de dados colunares utilizadas na DWT, assim como o método de configuração.

Valores de filtros passa baixa típicos dessa transformada são apresentados na Tabela 4.1. Utiliza-se o filtro Daub 8 em ambas as decomposições nesse trabalho. Entretanto, cabe ressaltar que a MODWT é usualmente computada com o filtro LA(8), por esse ser menos assimétrico, minimiza alguns dos efeitos observados. Esse filtro não está disponível no pacote utilizado, e assim pretende-se em trabalho futuro introduzir esse filtro com as devidas normalizações nesse pacote.

Tabela 4.1 - Valores não-zero dos filtros passa-baixa de Daubechies.

Daub 8		LA 8	
$k$	$h(k)$	$k$	$h(k)$
0	0.0544158422431072	4	0.069490465911
1	0.3128715909143166	5	-0.038493521263
2	0.6756307362973195	6	-0.073462508761
3	0.5853546836542159	7	0.0515398670374
4	-0.0158291052563823	8	1.099106630537
5	-0.2840155429615824	9	0.680745347190
6	0.0004724845739124	10	-0.086653615406
7	0.1287474266204893	11	-0.202648655286
8	-0.0173693010018090	12	0.010758611751
9	-0.0440882539307971	13	0.044823623042
10	0.0139810279174001	14	-0.000766690896
11	0.0087460940474064	15	-0.004783458512
12	-0.0048703529934520		
13	-0.0003917403733770		
14	0.0006754494064506		
15	-0.0001174767841248		

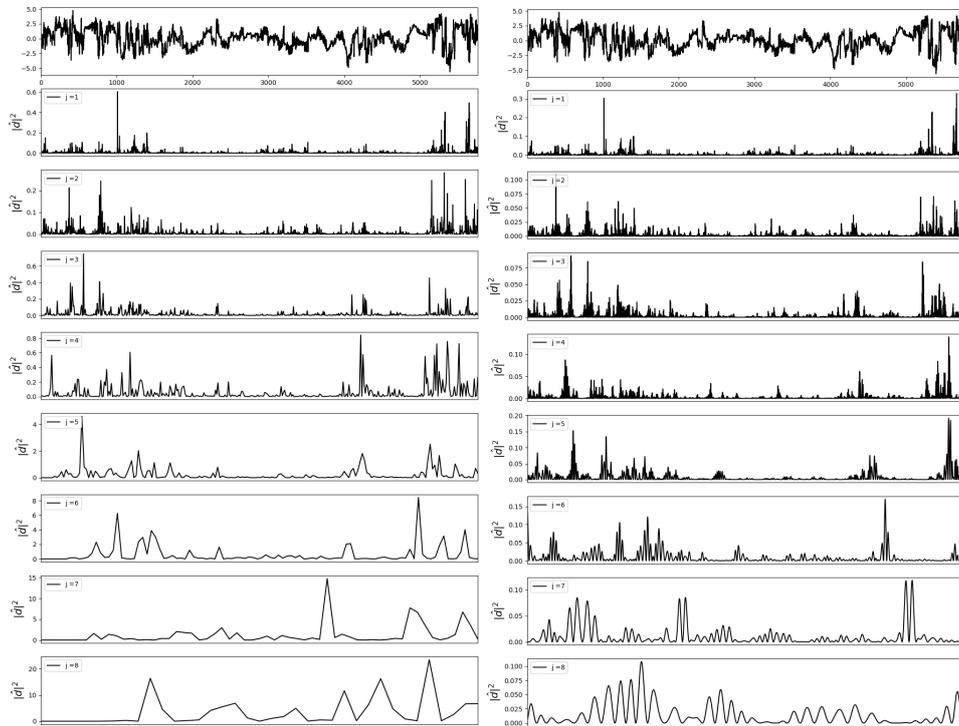
Fonte: Daubechies (1990).

#### 4.2.2 Resultados dos testes de sensibilidade

A visualização das contribuições das em multi-nível da DWT e MODWT é apresentada pelas energias obtidas dos módulos quadrados dos coeficientes *wavelets* ( $|d_k^j|^2$ ).

a) Como um primeiro estudo, considera-se a análise da decomposição em multi-nível.

Figura 4.3 - Decomposição multi-nível da componente  $B_z$  do Campo Magnético Interplanetário: DWT (esquerda) e MODWT (direita).

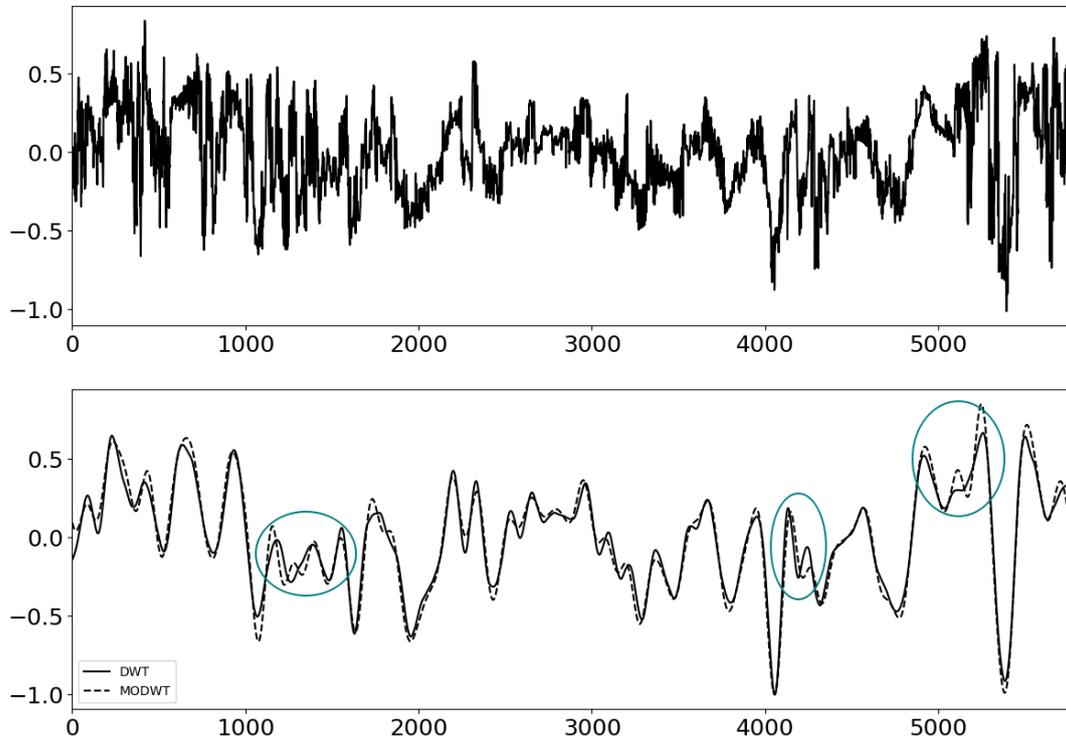


Fonte: Produção do Autor.

As duas decomposições apresentam diferenças principalmente nos níveis mais relacionados as baixas frequências, que devido a inexistência do processo de decimação na MODWT, torna-se notório a influência da redundância na decomposição nesses níveis.

b) Em uma segunda análise avalia-se a filtragem por escalas discretas. Devido a redundância e custo computacional da CWT, torne-se viável a utilização da DWT para realizar filtragem e extração de componentes frequenciais de séries temporais. Neste exemplo utiliza-se as decomposições em escalas apresentadas em 4.3 para realizar o processo de filtragem por escalas. A filtragem dos dados é dada pela extração dos níveis  $J = 5, 6, 7$  e  $8$  com a DWT ortogonal e a MODWT, apresentadas na Figura 4.4

Figura 4.4 - Série Temporal  $B_z$  (I) e , Filtragens com a DWT (Linha Contínua) e MODWT (Linha Tracejada) (II).



Fonte: Produção do Autor.

Observa-se que as séries reconstruídas apresentam maiores diferenças em certas regiões onde ocorreram grandes variações do sinal, como as indicadas na Figura 4.4. Isso ocorre pois a MODWT é redundante, nestes níveis estão também contidas informações de  $J = 4$  e  $9$  por exemplo. Isso permite que a MODWT retenha algumas informações não retidas na DWT devido a decimação. Entretanto, de forma geral, tem-se reconstruções similares conforme esperado mas com o mesmo número de dados em todas as escala. Essa situação favorece estudos em que seja necessário a comparação por escala dos dados com um mesmo número de elementos.

Em resumo, ao analisar o comportamento do sinal, pode-se utilizar a DWT sem prejuízos, mesmo ela sendo ligeiramente sensível a deslocamentos no sinal, porém, se for necessário realizar estudo de eventos com deslocamentos substanciais, recomenda-se utilizar a MODWT na filtragem, mesmo esta sendo mais cara computacionalmente.

## 5 ANÁLISE POR MÉTRICAS DE SIMILARIDADE

Um dos fundamentos importantes ao analisar um grande volume de séries temporais no contexto de mineração de dados, é a busca por similaridade entre dois ou mais conjuntos de dados (FU, 2011; RAFIEI; MENDELZON, 1997). Neste capítulo, como a terceira e última parte para estabelecer a metodologia integrada de análise de sinais, introduzem-se e estudam técnicas de similaridade como a da distorção temporal dinâmica (DTW, *Dynamic Timing Warping*), considerando suas características, de forma a prover métricas de avaliação de distâncias temporais. Estabelecem-se e demonstram-se, igualmente, para a técnica algumas de suas potenciais aplicações.

A similaridade de séries temporais pode ser determinada por meio de suas estruturas e formatos (FALOUTSOS et al., 1994). Uma métrica muito utilizada no estudo de similaridade em modelos é a *root-mean-square error* (RMSE), dada pela raiz do quadrado da diferença entre os elementos das duas séries, normalizadas pelo número de pontos, gerando um valor de erro médio entre os dados modelados e observados. Observa-se a formulação do RMSE em Equação 5.1, apresentada em (BARNSTON, 1992), sendo  $f_1$  uma série temporal observada e  $f_m$  uma série temporal modelada, então,

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (f_{1i} - f_{mi})^2}{N}}. \quad (5.1)$$

Outro método muito popular por sua simplicidade e eficiência é o cálculo da distância Euclidiana entre duas séries apresentada a seguir.

### 5.1 Distância Euclidiana

Dadas duas séries temporais  $f(t)$  e  $g(t)$ , em que:

$$f(t) = f_1, f_2, \dots, f_i, \dots, f_n \quad (5.2)$$

$$g(t) = g_1, g_2, \dots, g_j, \dots, g_m, \quad (5.3)$$

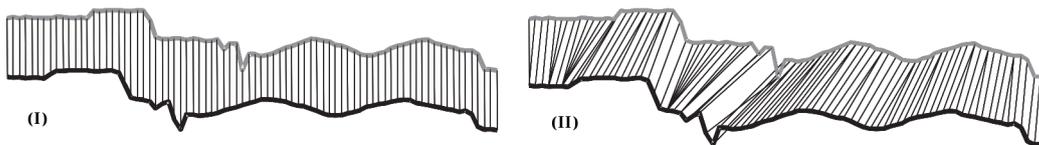
a distância euclidiana ( $\mathcal{D}$ ) é dada por:

$$\mathcal{D}(f(t), g(t)) = \sqrt{\sum_{i=1}^n (f_i - g_i)^2}. \quad (5.4)$$

Mesmo sendo o método mais popular dentro da comunidade de Mineração de Dados, a Distância Euclidiana apresenta pontos negativos. Um dos principais problemas a ser mencionado ocorre quando há deslocamento temporal em eventos correlacionados, a distância entre as duas séries será maior, pois a comparação não leva em conta deslocamento temporal de eventos (LIN et al., 2012). Para lidar com esse problema, outros métodos de análise foram criados. Uma alternativa para encontrar o alinhamento otimizado entre duas séries temporais contendo deslocamento em eventos, é a técnica *Dynamic Time Warping* (DTW) (MULLER, 2007; KEOGH; RATANAMAHATANA, 2005).

A Figura 5.1 apresenta uma ilustração da comparação ponto a ponto entre a Distância Euclidiana e o método DTW.

Figura 5.1 - Comparativo entre Distância Euclidiana (I) e a técnica *Dynamic Time Warping* (II).



Fonte: Adaptado de Keogh and Ratanamahatana (2005).

## 5.2 Dynamic Time Warping - DTW

Introduzida inicialmente por Sakoe (1978) a técnica DTW foi desenvolvida para detecção de padrões na fala humana. A proposta inicial surgiu na aplicação da técnica na transformação das ondas sonoras em séries temporais discretas.

A distancia DTW ( $\mathfrak{D}$ ) de  $f(t)$  e  $g(t)$ , é dada pela soma dos valores de  $\mathfrak{W}$ , que representa o caminho com o menor valor entre os valores da Matriz de Distâncias ( $\mathfrak{M}$ ). A equação 5.5 apresenta a computação necessária para obtenção dos valores das células da Matriz de Distância (BERNDT; CLIFFORD, 1994).

$$\mathfrak{M}_{i,j} = |f(t_i) - g(t_j)| + \min(\mathfrak{M}_{i-1,j-1}, \mathfrak{M}_{i-1,j}, \mathfrak{M}_{i,j-1}) \quad (5.5)$$

Para encontrar  $w_i$ , inicia-se pela célula  $\mathfrak{M}_{n,m}$  e soma-se os valores das células de menor valor vizinhas a célula de referência. O caminho  $\mathfrak{W}$  é dado por:

$$\mathfrak{W} = w_1, w_2, \dots, w_i, \dots, w_n, \quad (5.6)$$

O processo de exploração de séries temporais com grande volume de dados apresenta-se custosa, devido a complexidade dos muitos algoritmos existentes, para lidar com os gargalos de tempo das execuções dos algoritmos, métodos de simplificação de dados são criados (LIN et al., 2012). Keogh et al. (2001) descreve a utilização da Transformada Wavelet Discreta na redução de complexidade de séries temporais, utilizando processos de decomposição em escalas.

### 5.3 Testes de sensibilidade

Nesta seção considera-se alguns cenários associados a séries temporais formadas por funções seno com variações frequenciais em diferentes localizações temporais para analisar a resposta da ferramenta DTW e da distância Euclidiana.

São consideradas as seguintes séries temporais:

$$s_1(t) = \text{sen}(2\pi t), \quad s_2(t) = \text{sen}(10\pi t), \quad s_3(t) = \cos(2\pi t), \quad (5.7)$$

$$s_4(t) = \begin{cases} \text{sen}(2\pi t) & t \geq \pi \\ \text{sen}(10\pi t) & \text{caso contrário} \end{cases}, \quad s_5(t) = \begin{cases} \text{sen}(2\pi t) & t < \pi \\ \text{sen}(10\pi t) & \text{caso contrário} \end{cases}, \quad (5.8)$$

$$s_6(t) = \begin{cases} \text{sen}(2\pi t) & t \geq \frac{\pi}{2} \text{ e } t < \frac{3\pi}{2} \\ \text{sen}(10\pi t) & \text{caso contrário} \end{cases}, \quad s_7(t) = \begin{cases} \text{sen}(12\pi t) & t \geq \frac{\pi}{2} \text{ e } t < \frac{\pi}{2} + 0.8 \\ \text{sen}(2\pi t) & \text{caso contrário} \end{cases}, \quad (5.9)$$

$$s_8(t) = \begin{cases} \text{sen}(12\pi t) & t \geq \frac{3\pi}{2} \text{ e } t < \frac{3\pi}{2} + 0.8 \\ \text{sen}(2\pi t) & \text{caso contrário} \end{cases} \quad (5.10)$$

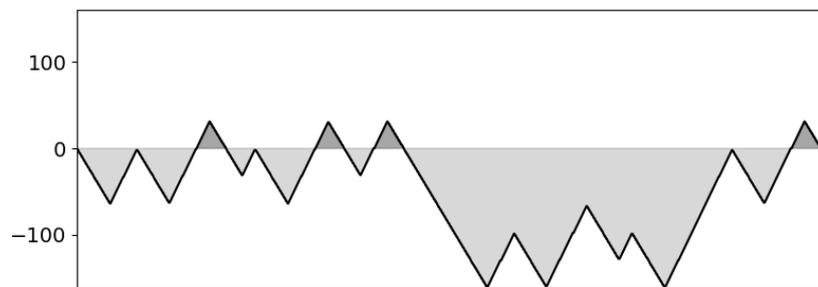
### 5.3.1 Estudo sobre o *Warping Path*

O caminho  $\mathfrak{W}$  (*warping path*) apresenta o alinhamento entre duas séries temporais obtidos pela DTW, tendo a continuidade temporal das séries como principal *constraint* (KEOGH; RATANAMAHATANA, 2005).

Neste teste de sensibilidade transforma-se o Caminho  $\mathfrak{W}$  em uma série temporal para que seja possível realizar observações sobre o seu comportamento. Utiliza-se o exemplo  $s_1 \times s_2$ , apresentado posteriormente na Figura 5.3 para ilustrar este estudo.

A Figura 5.2 apresenta o caminho  $\mathfrak{W}$  de  $s_1 \times s_2$  em formato de série temporal.

Figura 5.2 - Caminho  $\mathfrak{W}$ .



Fonte: Produção do Autor.

Observa-se que as amplitudes desta série temporal representa o quanto cada ponto deste caminho distancia-se da diagonal principal de  $\mathfrak{M}_{i,j}$ , os valores positivos desta série temporal, representam o quão adiantados os ajustes da segunda série temporal estão em relação a primeira, a lógica inversa se aplica também aos atrasos.

Uma análise que pode ser extraída do estudo do *warping path* no contexto de aferição de resultados de modelo, seria a integral das áreas de atrasos (área em cinza claro) e a integral da área de adiantamentos (área em cinza escuro), interpreta-se como o quanto o  $\mathfrak{w}$  se distanciou da diagonal principal em contraponto com a distância obtida no cálculo da DTW.

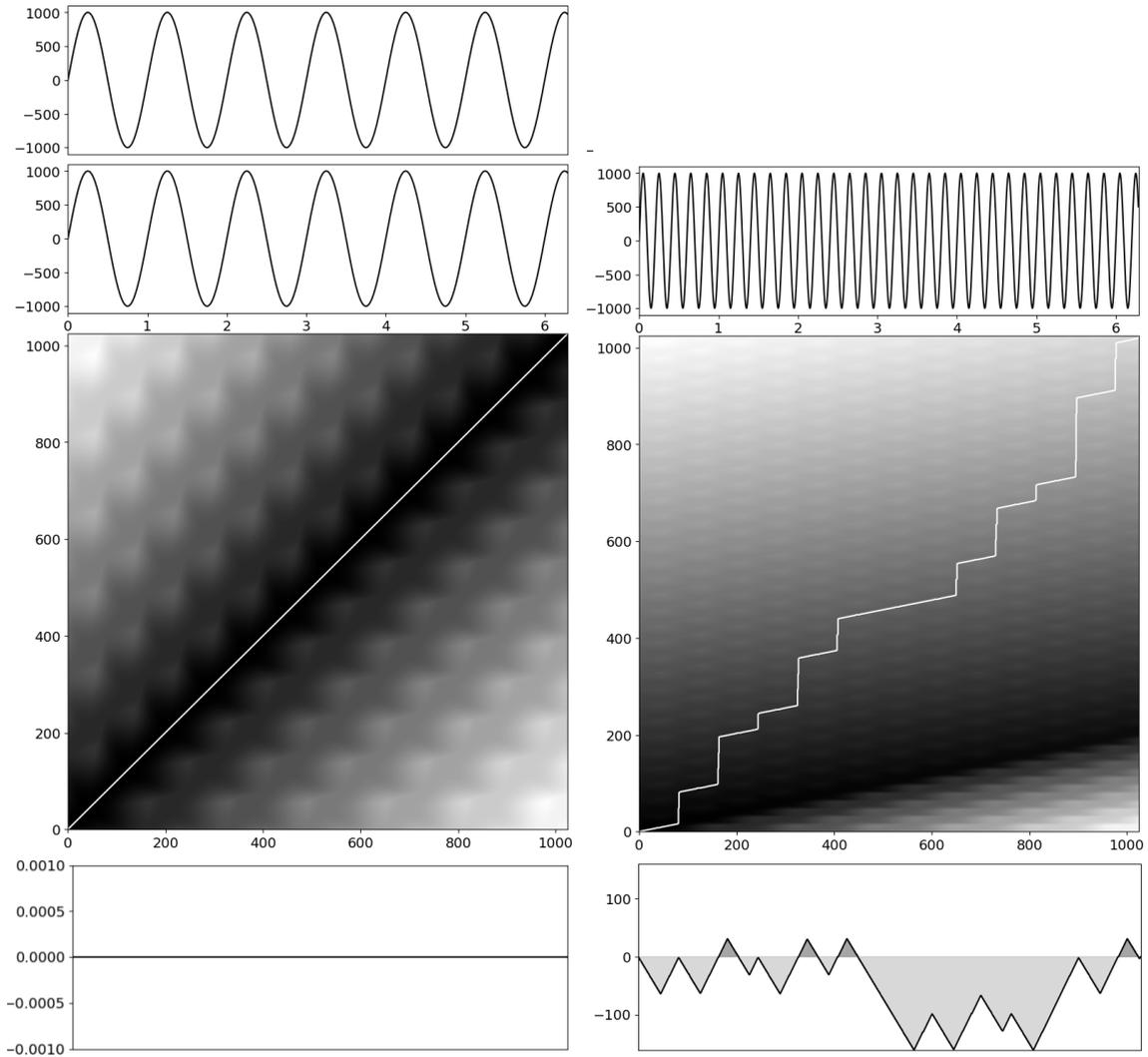
### 5.3.2 Cenário 1: diferentes frequências

O Cenário 1 apresenta testes das ferramentas distância euclidiana e DTW aplicadas a séries temporais  $s_1$  e  $s_2$ . Busca-se avaliar o comportamento dos indicadores globais

e da curva  $\mathfrak{W}$  ao se deparar com dados em diferentes frequências.

A Figura 5.3 apresenta os resultados dos testes do Cenário 1, os painéis a esquerda apresenta  $s_1 \times s_1$  e da direita  $s_1 \times s_2$ .

Figura 5.3 - Cenário 1: Séries Temporais, Matriz de Distância e Caminho  $\mathfrak{W}$ .



Fonte: Produção do Autor.

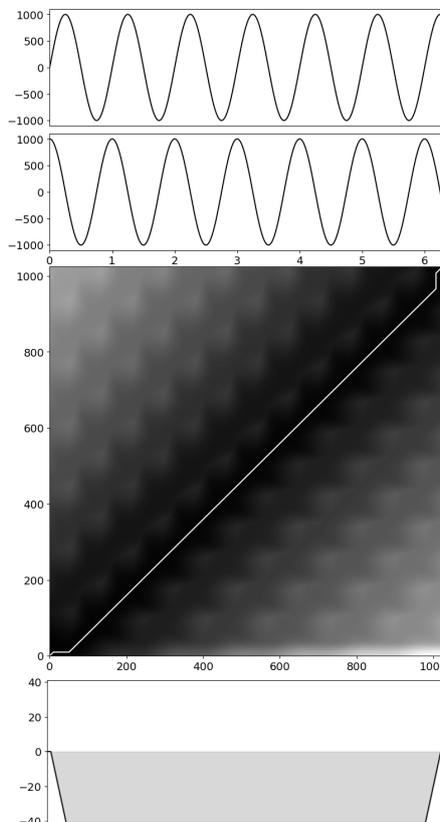
No exemplo  $s_1 \times s_1$  compara-se duas séries idênticas, um cenário em que ambas as ferramentas apresentam distância 0, nota-se que em  $\mathfrak{M}_{i,j}$  o caminho  $\mathfrak{w}$  é dado pela diagonal principal.

Comparando  $s_1 \times s_2$ , em que  $\mathcal{D} = 31,906$  e  $\mathcal{Q} = 0,260$  nota-se que a distância euclidiana apresenta um alto valor. Por meio de  $\mathfrak{w}$  observa-se que há uma uniformidade nos encaixes dos picos e vales de  $s_2$ , formando um caminho de baixa distância.

### 5.3.3 Cenário 2: diferentes fases

No cenário 2 realiza-se a comparação de  $s_1$  e  $s_3$ , que representa um cenário onde o segundo sinal possui um deslocamento de fase em relação ao primeiro. A Figura 5.4 apresenta o resultado do teste deste cenário.

Figura 5.4 - Cenário 2: Séries Temporais, Matriz de Distância e Caminho  $\mathfrak{W}$ .



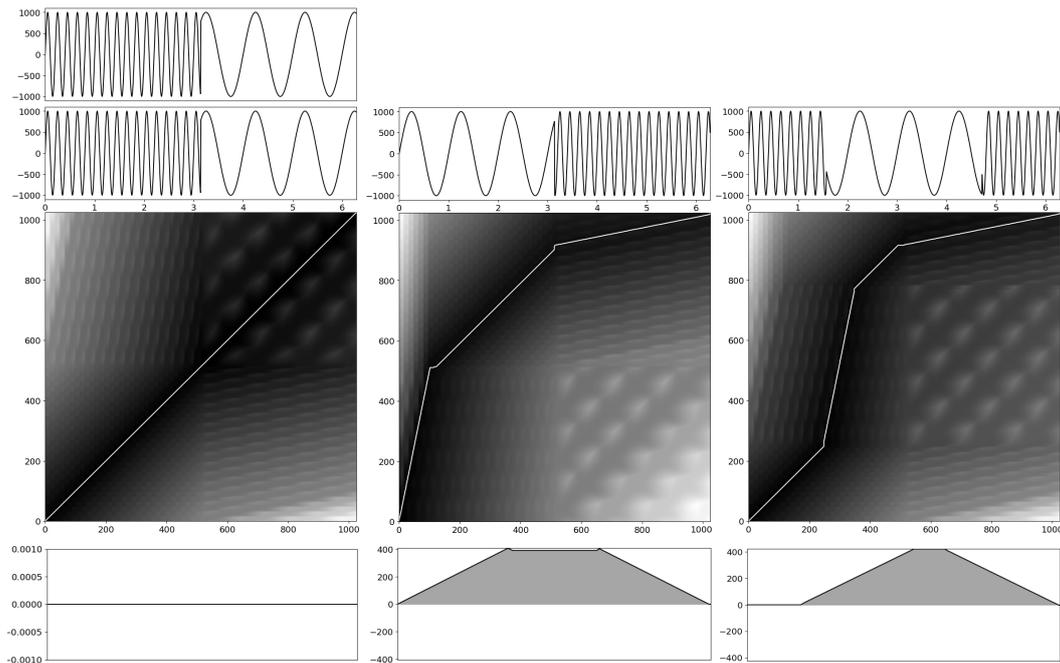
Fonte: Produção do Autor.

Por meio dos resultados globais deste cenário ( $\mathcal{D} = 32,527$  e  $\mathcal{Q} = 0,020$ ) e por  $\mathfrak{W}$  que a DTW detectou esse atraso de fase e alinhou todos os pontos seguintes considerando esse atraso, seguindo paralelo a diagonal principal.

### 5.3.4 Cenário 3: deslocamento temporal

Neste cenário apresentam-se séries temporais compostas por duas frequências em diferentes localizações temporais. A Figura 5.5 apresenta o resultado do teste deste cenário.

Figura 5.5 - Cenário 3: Séries Temporais, Matriz de Distância e Caminho  $\mathfrak{W}$ .



Fonte: Produção do Autor.

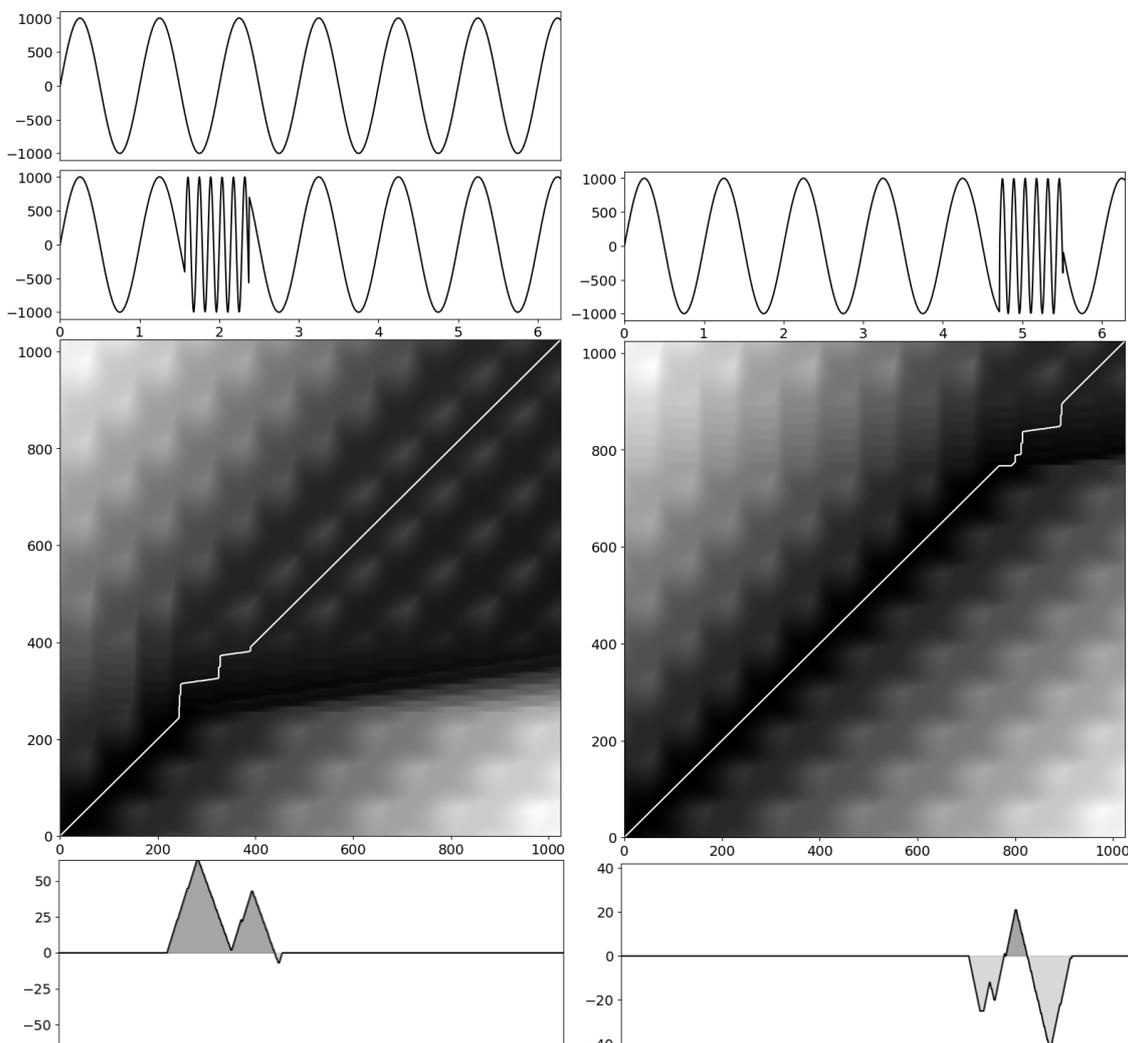
Neste cenário evidencia-se com clareza as diferenças entre os dois métodos estudados. No segundo e terceiro teste deste cenário obtém-se valores muito diferentes em distância euclidiana, pois no terceiro teste há o encaixe temporal entre as duas séries temporais durante o período inicial.

O comportamento de ajuste temporal da DTW fica visível nos  $\mathfrak{W}$  deste cenário, onde há um grande adiantamento de tempo entre as séries para a geração de  $\mathfrak{D}$ .

### 5.3.5 Cenário 4: localização de eventos

Neste cenário estudam-se séries temporais com eventos de oscilações rápidas localizadas (do inglês *burst*) em diferentes momentos de uma série temporal. A Figura 5.6 apresenta o resultado destes testes de sensibilidade.

Figura 5.6 - Cenário 4: Séries Temporais, Matriz de Distância e Caminho  $\mathfrak{W}$ .



Fonte: Produção do Autor.

Por meio dos resultados obtidos nos dois testes deste cenário (Tabela 5.1) nota-se que as duas técnicas não são sensíveis a eventos locais nos resultados globais dos cálculos de distância. A pequena diferença de valores entre os dois testes se dá a um pequeno atraso de fase gerado pelo evento, visto que no primeiro teste o evento acontece em um período anterior ao do segundo. A forma de identificar esse evento nestas séries temporais foram as perturbações apresentadas em  $\mathfrak{W}$ .

### 5.3.6 Resumo e comparação dos cenários

Conclui-se por meio dos testes de sensibilidade que os métodos de estudo de similaridade em séries temporais DTW e distância Euclidiana fornecem resultados globais, com baixa sensibilidade a localização de eventos. É possível observar também que pequenos deslocamentos de fase e mudanças frequenciais mudam totalmente os valores de  $\mathcal{D}$ , devido a característica de ajuste temporal dinâmico da DTW,  $\mathfrak{D}$  são menos sensíveis a essas mudanças. Observa-se também que  $\mathfrak{W}$  pode ser utilizado como uma ferramenta auxiliar visual para identificação dessas diferenciações entre os sinais.

A Tabela 5.1 apresenta o compilado dos resultados dos testes de sensibilidade.

Tabela 5.1 - Comparativo de distâncias Euclidiana e DTW.

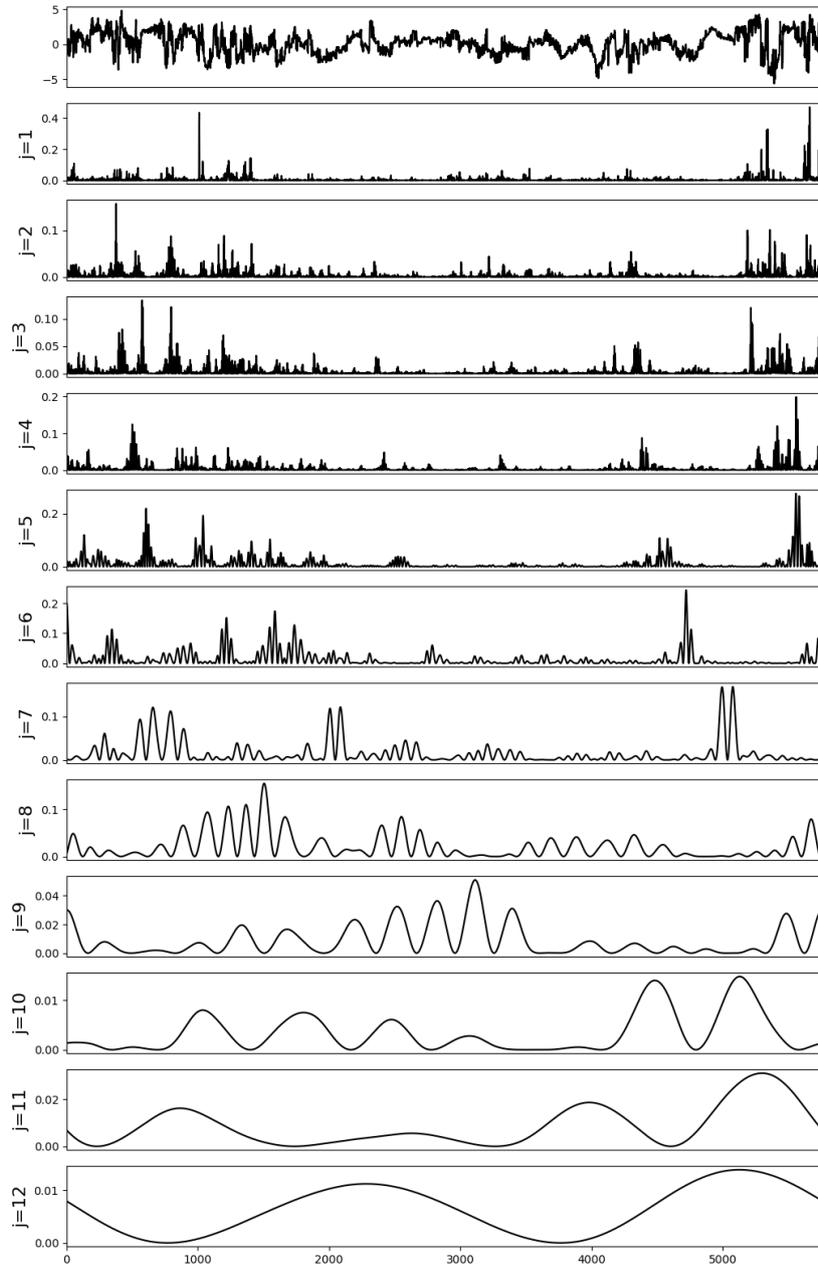
Cenário	Séries	Distância	
		Euclidiana ( $\mathcal{D}$ )	DTW ( $\mathfrak{D}$ )
$\neq$ Frequências	$s_1 \times s_1$	0,000	0,000
	$s_1 \times s_2$	32,566	0,271
$\neq$ Fases	$s_1 \times s_3$	32,527	0,020
Deslocamento	$s_4 \times s_4$	0,000	0,000
	$s_4 \times s_5$	32,570	0,027
	$s_4 \times s_6$	22,516	0,032
Localização de Eventos	$s_1 \times s_7$	12,277	0,043
	$s_1 \times s_8$	12,055	0,044

Fonte: Produção do Autor.

## 5.4 Resultados parciais: estudo de uma componente do campo magnético interplanetário

Nesta aplicação busca-se avaliar o comportamento da DTW e distância euclidiana, associando uma série temporal a sinais derivados, compostos por suas diversas escalas. Utiliza-se a componente  $B_z$  do campo magnético interplanetário, dados obtidos no período de 01/02/2018 a 04/02/2018. Para extrair três variações desta série temporal com diferentes filtragens, utiliza-se a decomposição em escalas em 12 níveis com a MODWT. A Figura 5.7 apresenta a decomposição em escalas de  $B_z$  com o filtro *Daubechies8*.

Figura 5.7 - Decomposição em escalas -  $B_z$ .



Fonte: Produção do Autor.

Nesta decomposição observa-se que as maiores contribuições de energia deste sinal estão contidas nas altas e médias frequências, as baixas frequências possuem baixas amplitudes. As séries temporais derivadas de  $B_z$  são compostas por um conjunto de escalas contendo variações de diversas origens físicas. A tabela 5.2 apresenta a composição das séries temporais utilizadas nesta aplicação.

Tabela 5.2 - Composição de escalas das séries temporais.

Séries	Níveis
$\ell B_z$	$j = 10, 11, 12$
$mB_z$	$j = 6, 7, 8, 9$
$hB_z$	$j = 1, 2, 3, 4$

Fonte: Produção do Autor.

Em  $\ell B_z$  estão contidas as componentes de baixas frequências (do inglês *low-frequency*, variações de 17 – 64 horas), em  $mB_z$  estão contidas as componentes de médias frequências (variações de 1 – 8 horas) e em  $hB_z$  estão contidas as componentes de alta frequência (do inglês *high-frequency*, variações de 2 – 32 minutos).

### Pré-Tratamento

Para analisar as variações das séries temporais normaliza-se os dados subtraindo a média aritmética da série e divide-se pelo máximo do módulo da série temporal, desta forma obtém-se séries temporais com valores entre  $-1$  e  $1$ .

Sendo  $B_z$  uma série temporal, a sua versão normalizada  $\widehat{B}_z$  é dada por:

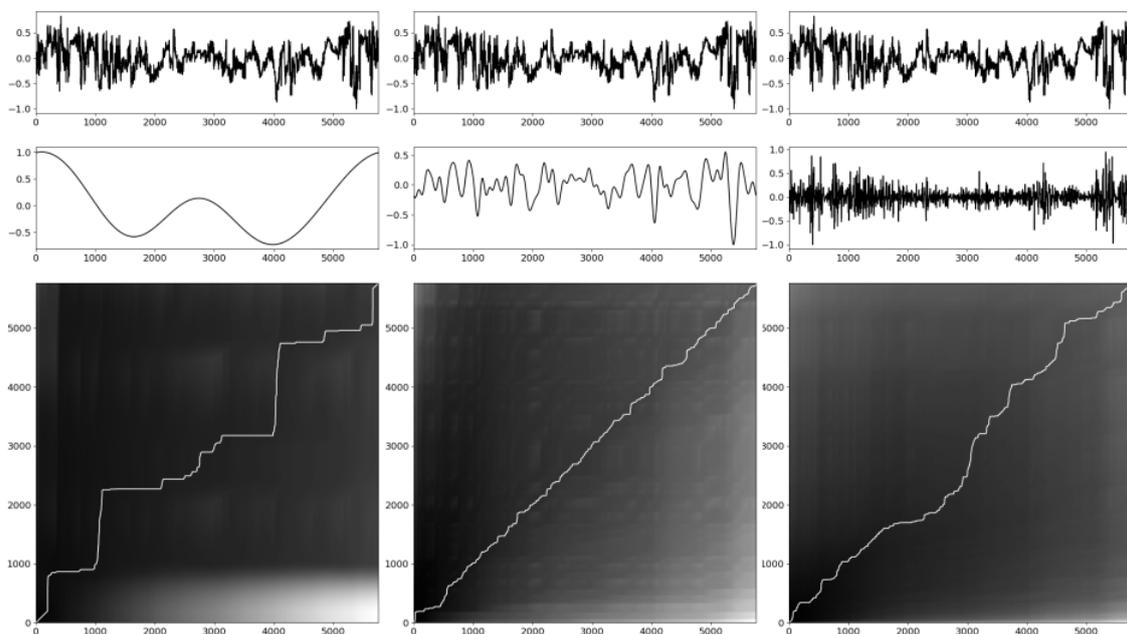
$$\widehat{B}_z = \frac{B_z - \overline{B}_z}{\max(|B_z|)}, \quad (5.11)$$

$\overline{B}_z$  é a média aritmética de  $B_z$ .

### Resultado

A Figura 5.8 apresenta os resultados da aplicação.

Figura 5.8 - Aplicação: Séries Temporais e Matriz de Distância.



Fonte: Produção do Autor.

Na comparação  $B_z \times \ell B_z$ , com a perda dos detalhes de mais altas frequências, obteve-se um maior valor da distância euclidiana e DTW em relação as outras duas medições ( $\mathcal{D} = 37,755$ ,  $\mathfrak{D} = 0,115$ ). Observa-se pela decomposição em escalas na Figura 5.8 que as contribuições das baixas frequências são menores em relação as altas e médias, torna-se natural essa maior distância visto que o comportamento de baixa frequência não caracteriza bem o sinal original. Outro aspecto a ser observado, é a quantidade de adiantamentos e atrasos em  $\mathfrak{W}$ , e suas distâncias da diagonal principal, representando a quantidade de pontos que foram deslocados para esse encaixe temporal.

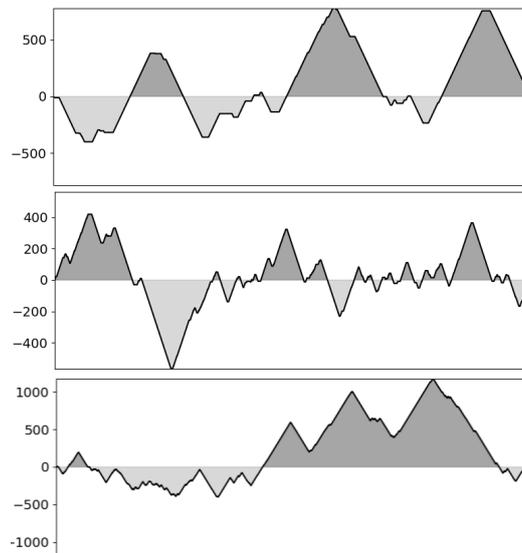
Nos quadros ao centro da Figura, é possível observar que  $mB_z$  adotou os comportamentos mais visíveis de  $B_z$ , isso também se traduziu também em  $\mathfrak{W}$ , que assumiu um comportamento mais próximo a diagonal principal, em que os principais desvios atrasos foram apresentados onde localmente as variações de alta frequências apresentam amplitudes predominantes.

Observa-se que em  $B_z \times mB_z$  apresenta-se  $\mathcal{D}$  menor do que em  $B_z \times hB_z$ , porém esse resultado se inverte em  $\mathfrak{D}$ . Esta inversão deve-se às inúmeras variações rápidas com ajuste em um mesmo ponto, conforme atrasos apresentados em  $\mathfrak{w}$  do quadro a

direita, gerando assim uma distância global mais baixa para a DTW.

Apresenta-se em 5.9 a representação de  $\mathfrak{W}$  das três análises.

Figura 5.9 - Aplicação - Caminho  $\mathfrak{W}$  1)  $B_z \times \ell B_z$ , 2)  $B_z \times \ell m B_z$ , 3)  $B_z \times h B_z$ .



Fonte: Produção do Autor.

Conforme esperado, a DTW de  $B_z \times \ell B_z$  apresenta maiores amplitudes no deslocamento da diagonal principal, justifica-se esse comportamento pela ausência de variações rápidas em  $\ell B_z$ , as variações médias do sinal possuem grande contribuição na formação da estrutura global, isso fica observável por meio da baixa amplitude dos deslocamentos do  $\mathfrak{W}$  na análise  $B_z \times m B_z$ . Já no caso de  $B_z \times h B_z$ , mesmo possuindo uma distância DTW menor, devido ao grande número de ajustes possíveis, proporcionados pelas variações de alta frequência, observa-se uma distância euclidiana maior em relação a  $m B_z$ . A Tabela 5.3 apresenta o compilado dos valores quantitativos obtidos nesta aplicação para a componente  $B_z$ .

Tabela 5.3 - Comparativo de distâncias Euclidiana e DTW.

Séries	Distância	
	Euclidiana	DTW
$B_z \times \ell B_z$	37,755	0,115
$B_z \times mB_z$	15,482	0,051
$B_z \times hB_z$	17,887	0,046

Fonte: Produção do Autor.

## 6 APLICAÇÕES, RESULTADOS E DISCUSSÕES

Neste Capítulo, aplicam-se as metodologias investigativas de padrões em sinais, buscando sobretudo características de completude ferramental de análise em uma área de fenômenos não-lineares e multiescala.

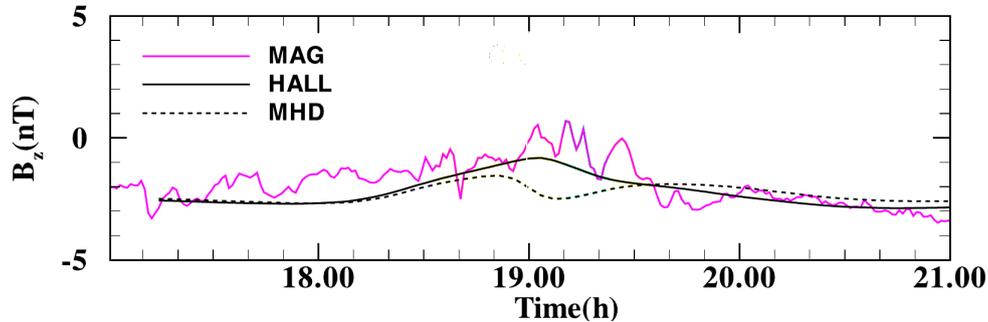
Em uma primeira parte, consideram-se a análise de dados observacionais coletados por sonda espacial e dados provenientes de simulações MHD relacionados ao mesmo processo físico. As simulações usam duas condições de modelo: a abordagem resistiva e a abordagem com o termo Hall. Em uma segunda parte, consideram-se a análise de um agente causador, controlado essencialmente pela componente  $B_z$  do campo magnético interplanetário, e as respostas físicas do sistema, explicitadas pelas mensurações da perturbação geomagnética, aqui os índices geomagnéticos  $AE$ , efeito do eletrojato auroral, e ASY-H, efeito de correntes elétricas relacionadas em baixa latitude. Ainda nesse caso, possibilitando uma análise adequada, empregam-se duas situações, uma em que há a condição de calma geomagnética e a outra, a condição de perturbação geomagnética.

Para cada uma das partes, procederam-se análises de relações tempo-escala entre dados e das métricas de similaridade entre dados, com o estabelecimento respectivo de interpretações.

### 6.1 Aplicação da metodologia a dados observacionais de satélite e dados de simulação

Com os uso dos dados observacionais da sonda Cassini, fazem-se as comparações das características dos sinais confrontando com as dos dados de simulações MHD resistivo (rMHD) e MHD Hall (hMHD). Os dados referem-se às observações e simulações no intervalo entre as 17:00 e 21:00h (tempo universal da Terra) em Dezembro de 2005. Como caso em análise nesta seção, a Figura 6.1 apresenta a componente  $B_z$  do campo magnético (MAG) medido pela sonda Cassini em trânsito perto de Titan, satélite de Saturno, e resultados das modelagens MHD resistivo (MHD) e com termo Hall (HALL). As séries possuem 1310 registros.

Figura 6.1 - Componente  $B_z$  do campo magnético (MAG) medido pela sonda Cassini em trânsito perto de Titan, satélite de Saturno, e resultados das modelagens MHD resistivo (MHD) e com termo Hall.



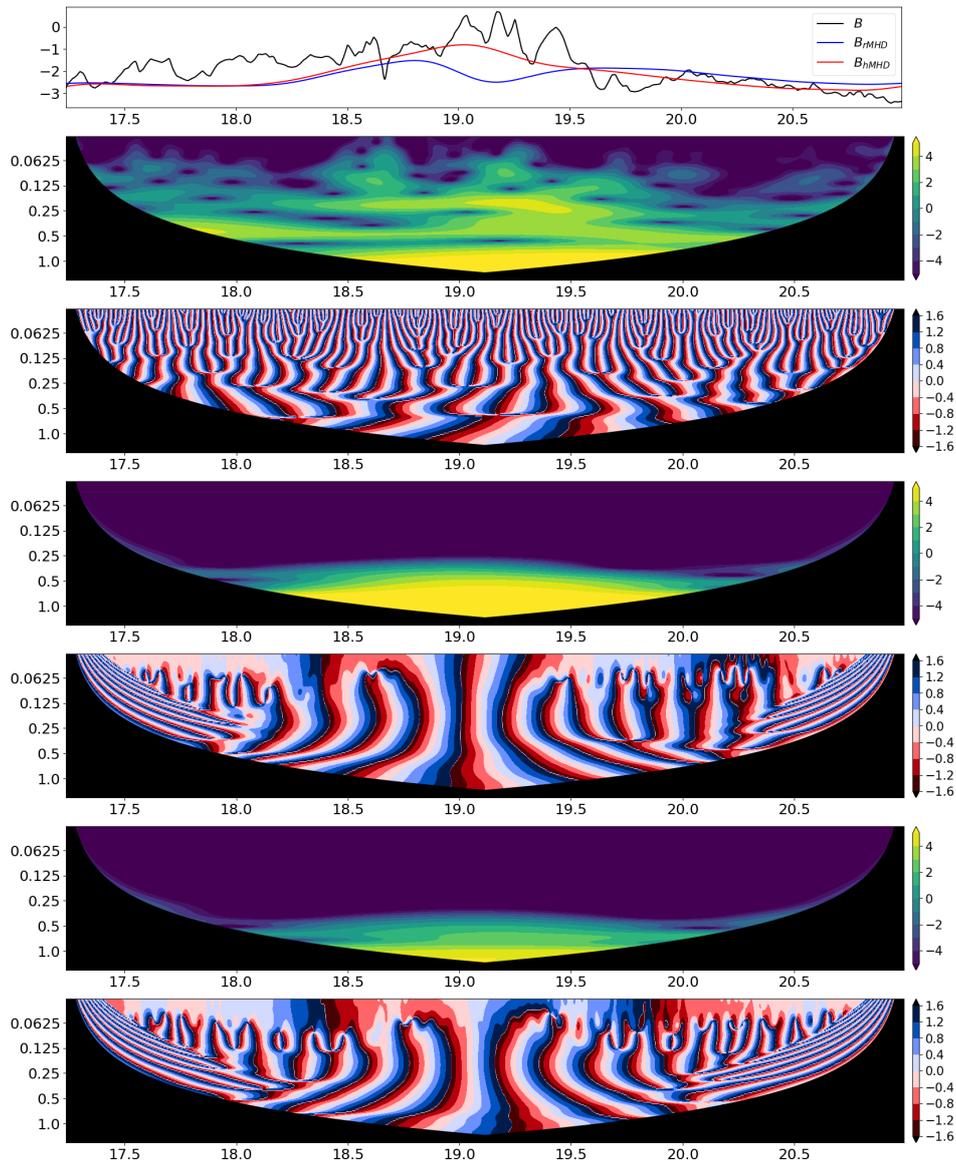
Fonte: Adaptado de Ma et al. (2007).

Na figura, por meio da interpretação por inspeção visual, embora constate-se certa similaridade de comportamento entre os dados e os resultados das simulações, há diferenças aparentes de valores e comportamentos entre os dados que ocorrem ao longo do intervalo. Desta forma, uma questão natural suscita do esforço de entendimento: como caracterizar melhor os sinais de maneira a poder revelar ou, no mínimo, auxiliar no entendimento da consistência de comportamento entre observações experimentais e resultados de aplicações de modelagem numérico-computacional?

### 6.1.1 Análise de relações tempo-escala

Por meio da técnica de análise tempo-escala baseada em transformadas *wavelets*, analisam-se as características dos dados da componente  $B_z$  do campo magnético e dos dados dos modelos rMHD e hMHD, que estão representadas na Figura 6.2.

Figura 6.2 - Dados da componente  $B_z$ : escalogramas e fases.



Fonte: Produção do Autor.

O primeiro painel, na parte superior, mostra os dados. No segundo painel, mostra-se o escalograma e, no terceiro painel, as fases por escala. No quarto e quinto painel, mostram-se respectivamente o escalograma e a fase por escala para os dados do modelo rMHD. E, por último, nos painéis, mostram-se de forma similar para os dados do modelo hMHD.

No escalograma do dado observacional, é possível verificar estruturas distribuídas em todas as escalas, o que revela a complexidade do sinal. Igualmente, no gráfico da

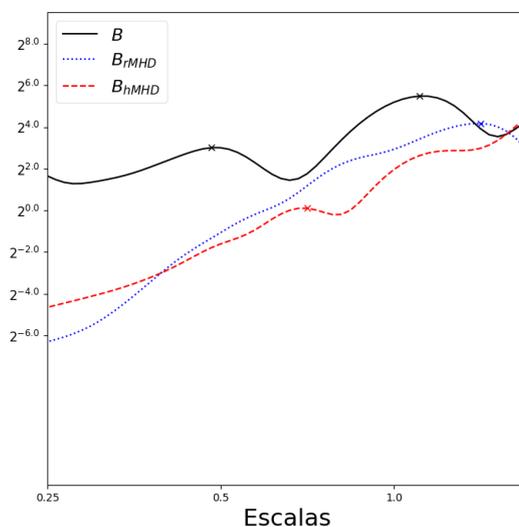
fase, observam-se mudanças complexas ao longo do tempo, quanto às escalas, nos padrões do sinal.

No escalograma dos dados do modelo rMHD, observam-se estruturas apenas entre as escalas de 0,25 a 1,5 hora (ou um quarto a uma hora e meia). Nota-se que os dados do modelo resistivo, embora trazendo características do sinal físico visto no escalograma das medidas, não consegue como simulação produzir respostas em escalas menores, isto é, em frequências mais altas. O gráfico da fase, pelo seu comportamento comparado ao das medidas, sugere alguma similaridade da fase do sinal, conquanto bastante difusa, isto é, muito geral nas indicações das fases.

Analisando o escalograma dos dados do modelo hMHD, percebe-se um comportamento similar à análise anterior. Constata-se um destaque um pouco diferente da contribuição de energia pelas escalas e ao longo do tempo; conquanto os aspectos gerais estejam algo preservados. O gráfico de fase, como o anterior, traz somente uma sugestão difusa de comportamento. No entanto, entre as fases dos sinais para rMHD e hMHD, há detalhes que distinguem relativamente os resultados das duas simulações.

De forma a possibilitar uma visão mais global, equivalente ao usual com outra técnica clássica de análise (ou seja, por meio das funções de Fourier), obtém-se da distribuição de energia por escala, representada na Figura 6.3, o espectro *wavelet* global para os dados observacionais  $B_z$  (em preto) e os respectivos espectros para os dados simulados rMHD (em azul pontilhado) e hMHD (em vermelho tracejado).

Figura 6.3 - Espectro *wavelet* global para componente  $B_z$  referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado).



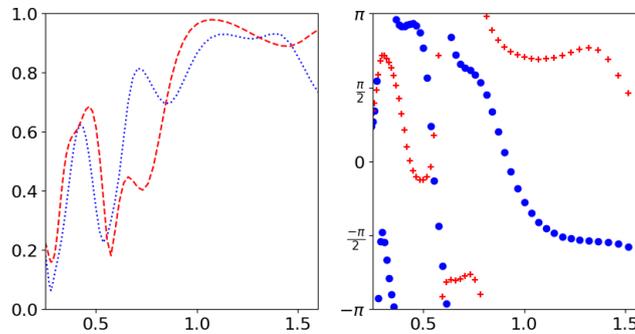
Fonte: Produção do Autor.

No espectro global da medida estão marcados os picos principais. As curvas dos valores simulados estão também mostradas, rMHD em azul pontilhado e hMHD em vermelho tracejado. As tendências de comportamentos são percebidas; embora as relações não sejam evidentes. Daí a importância de não se usar apenas um espectro global de sinais; porém análises do tipo tempo escala (escalograma) e, como visto adiante, outras abordagens *wavelet* derivadas.

Na Figura 6.4 estão apresentadas as correlações por escala (painel à esquerda), variando entre 0 e 1, e fase (à direita), variando entre  $-\pi$  e  $\pi$ , entre os dados observacionais de  $B_z$  e os obtidos pelos modelos rMHD (curva em azul pontilhado) e hMHD (curva em vermelho tracejado). No gráfico da correlação, notam-se alguns picos de correlação por escala. Entre as escalas 0,4 e 0,5h, há um pico referente a  $R \approx 0.7$  (ou  $R^2 \approx 50\%$ , que se traduz em termos de uma percentagem de flutuações correlacionadas entre dois sinais). Outro pico ocorre entre as escalas de 0,65 e 0,75h destacado para rMHD,  $R \approx 0.8$  (ou  $R^2 \approx 64\%$ ), e insinuado para hMHD,  $R \approx 0.4$  (ou  $R^2 \approx 16\%$ ). Um outro intervalo entre 1,0 a 1,25h apresenta um pico para rMHD,  $R \approx 0.9$  (ou  $R^2 \approx 81\%$ ), e para hMHD,  $R \approx 0.98$  (ou  $R^2 \approx 96\%$ ). Observam-se muito boa correlação entre  $B_z$  e os dados simulados rMHD e hMHD no intervalo de escala entre 1,0 a 1,25h. Há uma correlação também comum às duas simulações, embora atenuada em intensidade, em torno da escala de 0,5h. Em torno da escala

de 0,75h, há outra correlação, em que as proporções de contribuições para os sinais são distintas em termos das simulações executadas. Para o entendimento adequado, as análises das fases entre  $B_z$  e os dados das simulações referentes às escalas são mostradas no painel à direita. Para a escala 0,5h, os sinais estão praticamente em fase. Para a escala 0,75h, os sinais podem ser considerados aproximadamente em fase. E, finalmente, para as escalas entre 1,0 e 1,25h, os sinais não estão nem ajustados em fase nem em quadratura perfeita.

Figura 6.4 - Módulo e fase da correlação por escala da componente  $B_z$  em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado).

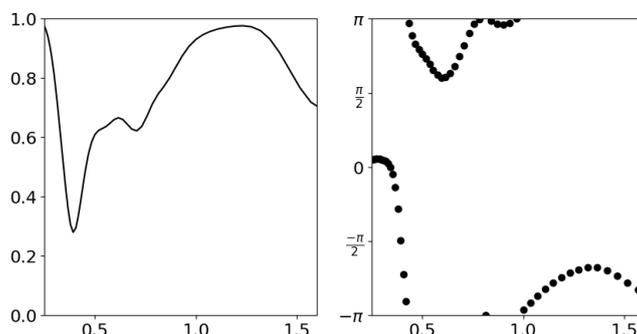


Fonte: Produção do Autor.

Estendendo os esforços de compreensão das medidas e dos dados simulados, a metodologia *wavelet* permite ainda a realização de outras análises significativas.

Na Figura 6.5 está apresentada a correlação por escala (módulo no painel à esquerda e fase, à direita) entre os dados simulados da variável  $B_z$  pelos modelos rMHD e hMHD.

Figura 6.5 - Módulo e fase da correlação por escala dos dados da componente  $B_z$  dos modelos rMHD e hMHD.

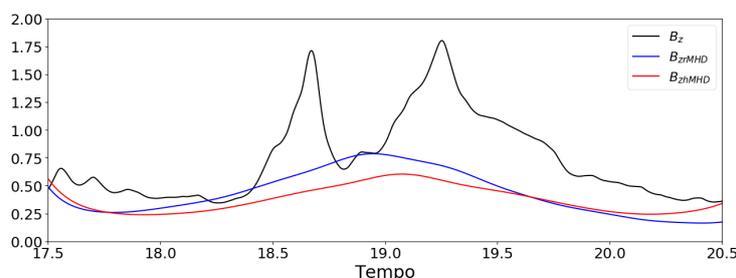


Fonte: Produção do Autor.

Para as escalas dos valores simulados (entre as escalas de 0,25 a 1,5h), a correlação começa no máximo valor, isto é,  $R^2 = 100\%$  de flutuações inter-relacionadas, diminui até um mínimo ( $R^2 \approx 10\%$ ) próximo a escala 0,5h, e volta a aumentar com pico próximo à escala 0,75h ( $R^2 \approx 50\%$ ), até atingir em 1,25h um valor elevado ( $R^2 \approx 96\%$ ), que é próximo ao máximo. No painel da direita, da análise de fase da correlação em termos da escala, nota-se que o sinal estava em fase, varia e volta a um comportamento em fase (em torno da escala de 0,5h), varia e retorna a uma situação de fase (em torno de 0,75h) e, finalmente, sai da fase e não alcança uma quadratura (em torno de 1,25h). As interpretações desta última figura e da figura 6.4 destacam assim as escalas principais entre os dados medidos e os simulados, como também as características dos sinais reconstruídos com essas escalas em comum.

Relativa à complexidade dos dados, pode-se fazer uma análise da entropia dos sinais em estudo, cujos resultados estão representados na Figura 6.6.

Figura 6.6 - Entropia *wavelet* da componente  $B_z$  e dos valores dos modelos MHD.

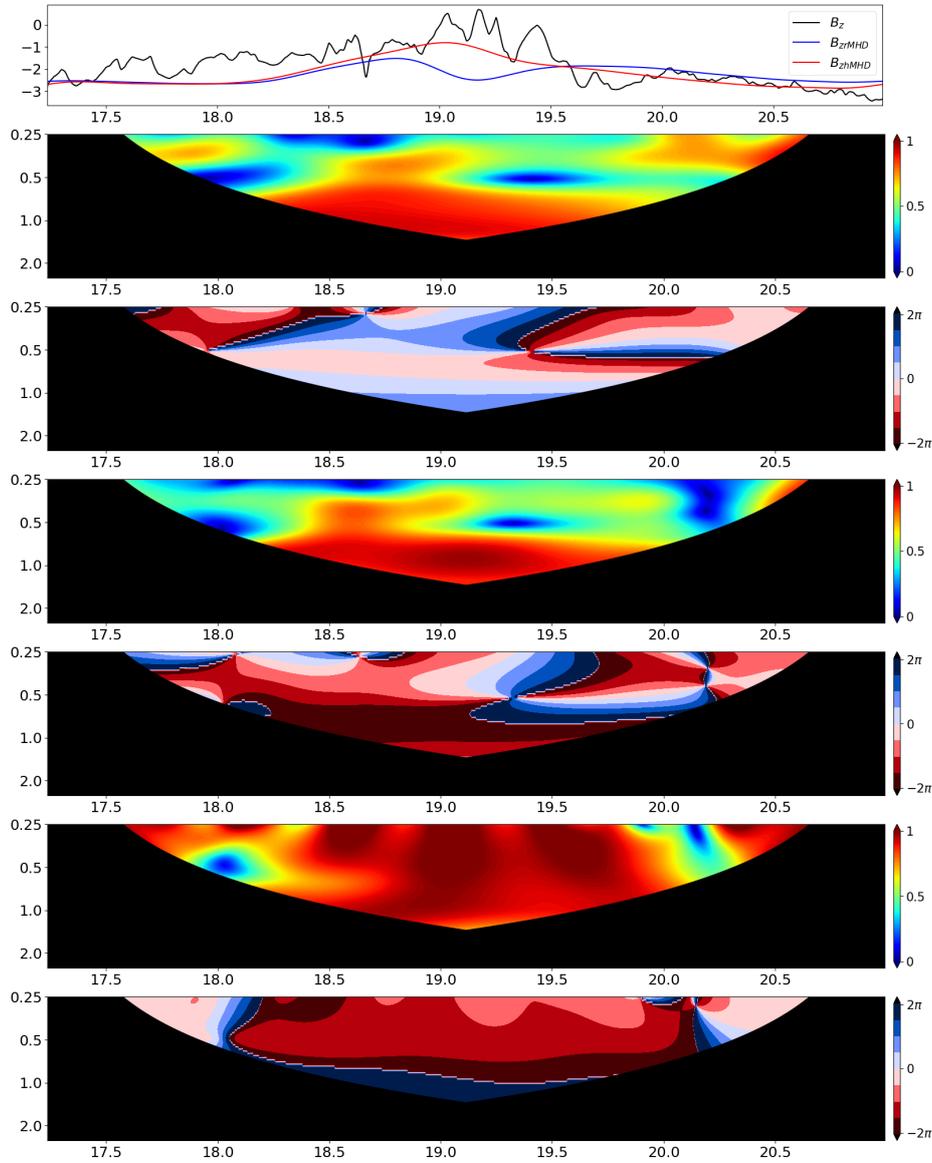


Fonte: Produção do Autor.

O comportamento da curva referente à  $B_z$  (em preto) destaca a complexidade existente nas medidas físicas, que, embora presente nos valores simulados (resistivo em azul pontilhado e Hall em vermelho tracejado) em um grau significativo, não tem condição de ser capturada na forma atual com que as simulações magneto-hidrodinâmicas foram realizadas. Essa caracterização não desmerece o valor das simulações na tentativa de descrever comportamentos físicos; no entanto, ela sinaliza que aprimoramentos dos modelos computacionais são necessários.

Considerando o intervalo efetivo de escalas (entre 0,25 e 1,5h) determinado dos modelos de simulação, calculam-se as análises de coerência baseada nas transformadas *wavelet* entre, inicialmente, as medidas  $B_z$  e os dados de simulação e, depois, os próprios dados da simulação, cujos resultados estão apresentados na Figura 6.8. No primeiro painel, mostram-se os sinais considerados nos estudos (a medida e os dados de simulação). De cima para baixo, mostram-se em pares os gráficos das análises de coerência *wavelet* e diferença de fases da coerência entre a medida e o dado rMHD, a medida e o dado hMHD e os dados rMHD e hMHD, respectivamente.

Figura 6.7 - Coerência *wavelet* e diferença de fases para a relações medida  $B_z$  e dado rMHD,  $B_z$  e dado hMHD e entre rMHD e hMHD.



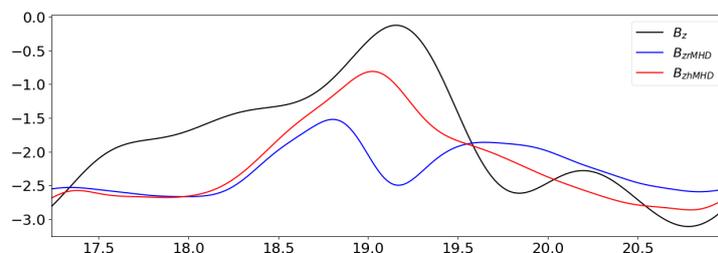
Fonte: Produção do Autor.

Da análise dos gráficos de coerência *wavelet*, constata-se explicitamente as relações de sinais por escala e por tempo, o que já tinha sido inferido da análise anterior da correlação por escalas. A cor quente (vermelha) indica as intensidades maiores da coerência e a cor fria (azul) as intensidades menores, havendo uma barra de cores (variando de 0 a 1) disponível à direita nos gráficos. Comparando os dois primeiros gráficos de coerência, observam-se as relações das medidas e dos dados das simulações MHD, em que se constata similaridade de comportamento. Alguns

detalhes distinguem as duas situações de simulações. Nos gráficos das diferenças de fase, observam-se as características dos sinais quanto às defasagens de fase por escala no tempo. Para a escala em torno de 0,5h, nota-se explicitamente a situação de fase do sinal. Em torno de 0,75h, constata-se também o comportamento em fase. Já para o intervalo 1,0 a 1,25h, conquanto afetado pela restrição de uma região de valor central, a relação de sinais apresenta um afastamento do acompanhamento em fase. Nesta última constatação, a percepção qualitativa é auxiliada pela avaliação quantitativa calculada na Figura 6.5. Referente aos gráficos da diferença de fase, os valores, codificados à direita na barra de cores, variam de  $-2\pi$  (cor vermelha) a  $2\pi$  (cor azul), constituindo representação de um círculo trigonométrico. Desta forma, por meio da análise de coerência para a relação dos dados dos modelos resistivo e Hall, nota-se a consistência das relações logo acima interpretadas. No entanto, as regiões em cor fria (tons de azul) relativas às escalas selecionadas denotam as particularidades onde a relação está enfraquecida ou inexistente.

Tais escalas predominantes, caracterizadas pelas correlações e comportamentos de coerência, permitem reconstruir metodologicamente os sinais, mostrados na Figura 6.8, e evidenciar de maneira mais definida as inter-relações existentes.

Figura 6.8 - Sinais reconstruídos, pela aplicação da técnica *wavelet*, para os dados das medidas e das simulações.



Fonte: Produção do Autor.

Vê-se que a curva da medida fica despojada de uma complicação inicial, que era uma maior complexidade no seu comportamento ao longo do tempo. Sua ocorrência de variação mostra os comportamentos básicos do sinal, neste caso com um crescimento, uma modulação, um intervalo de pico e um retorno a valor de fundo com flutuação. Já o resultado da modelagem resistiva (rMHD) tem um desenvolvimento mais suave, depois demonstra um pico na ocorrência e um retorno, seguido de uma nova elevação em uma curva suave que retorna a um valor de fundo. Enquanto o re-

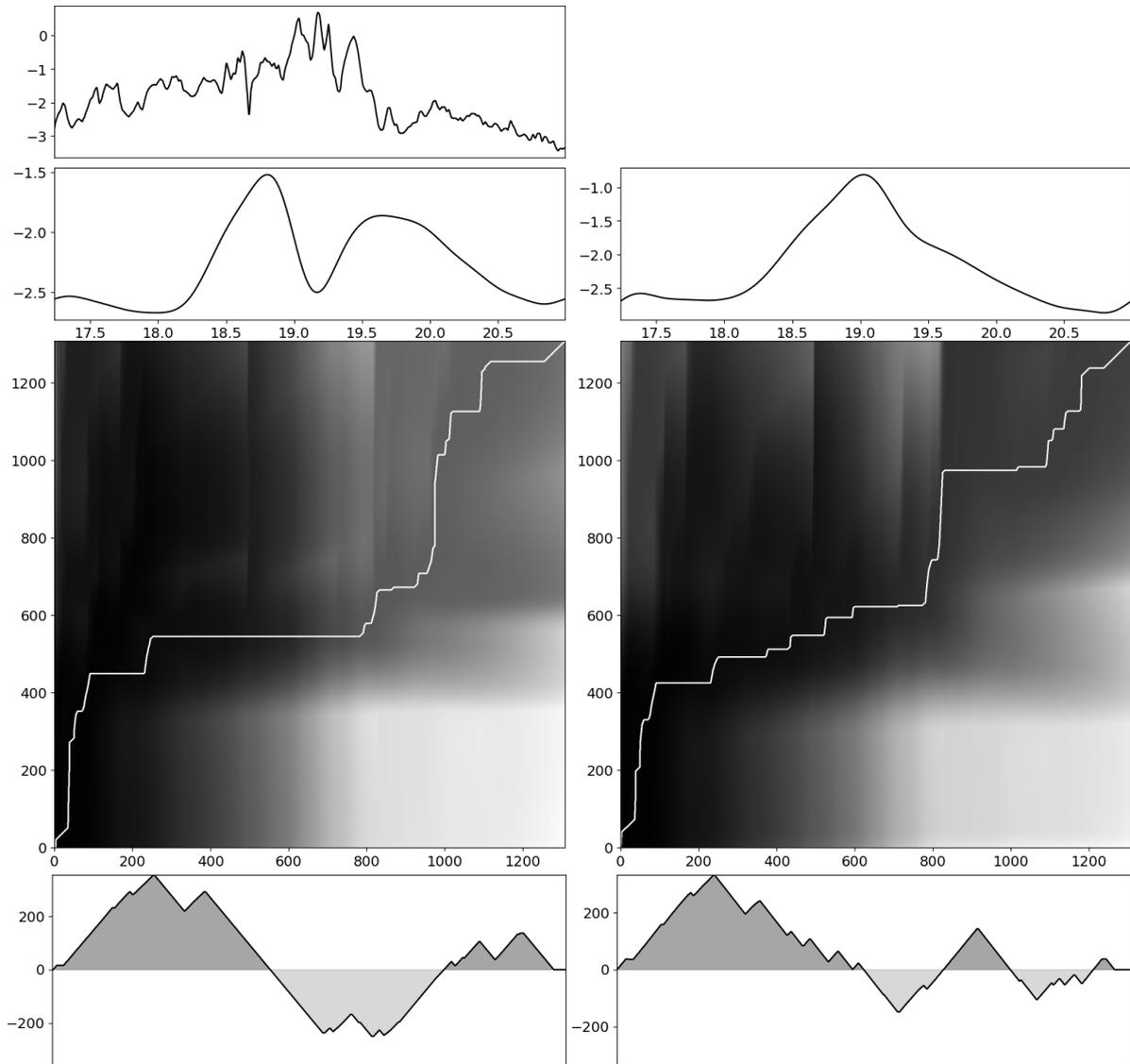
sultado da modelagem Hall (hMHD) tem um desenvolvimento suave, uma elevação mais significativa de valor aproximando do comportamento da medida, um decaimento que leve variação e um retorno ao valor de fundo, que de alguma forma é comportamento médio considerando os valores das medidas. Ao longo da investigação metodológica, de forma global, observa-se que os dados simulados pelos modelos magneto-hidrodinâmicos capturam algumas escalas de forma mais destacadas em relação a todas as escalas presentes nas medidas, no caso as obtidas pela sonda espacial Cassini. Considerando as contribuições em energia por escala no sinal ao longo do tempo, há similaridade entre os resultados dos modelos rMHD e hMHD. No entanto, notam-se algumas diferenças que se mostram como estruturas localizadas. Não está no escopo deste trabalho explorar fisicamente os fenômenos, embora a metodologia demonstre o seu valor por possibilitar explicitações sejam tanto em caráter geral e comum quanto em aspectos localizados e particulares.

### **6.1.2 Análise das métricas de similaridade entre dados observacionais e dados simulados**

A representação da inter-relação dos dados alcançada na seção anterior, neste caso, medidas e dados de simulação, possibilita aplicar mais adequadamente uma avaliação de métrica de similaridade. Esta análise acrescenta capacidade de interpretação da evolução do fenômeno físico que se faz envolvendo complicados mecanismos do ambiente espacial.

Usando a técnica DTW, avaliam-se os deslocamentos relativos temporais entre os comportamentos na medidas e nos dados das simulações MHD, no caso quanto aos modelos resistivo (rMHD) e Hall (hMHD). A Figura 6.9 apresenta a aplicação da técnica e os seus resultados em termos gráficos.

Figura 6.9 - Comparativo de matrizes de distância e caminhos  $\mathfrak{W}$  entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita).



Fonte: Produção do Autor.

No painel superior encontra-se a medida obtida pelo satélite. Os dois painéis seguintes apresentam os dados para rMHD, à esquerda, e hMHD, à direita. Os painéis em tom de cinza apresentam, respectivamente, as relações de deslocamento de tempo entre a medida e os dados de simulação. A curva em branco representa o caminho que otimiza por meio de algoritmo a relação de tempo entre os elementos de cada sinal, que explicitaria a relação física mais provável entre os sinais. Essas mesmas curvas determinam o gráficos, nos painéis na parte mais inferior, em que o atraso do

sinal de resposta com respeito à medida está identificada na curva superior ao eixo (cinza escuro) e os retardos na parte inferior (cinza claro). O resultado permite complementar os resultados das análises *wavelets*. Embora sem entrar no contexto físico dos dados, com respeito aos aspectos de similaridade e dissimilaridade, acrescenta-se uma percepção de que o resultado do modelo hMHD, comparado ao rMHD, segue de forma mais próxima a o comportamento  $B_z$ .

Abordando quantitativamente a análise de similaridade, no primeiro cenário obteve-se área de 122445 pontos de atrasos e 223494 pontos de adiantamentos para o caminho  $\mathfrak{W}$ . No segundo cenário obteve-se área de 44202 pontos de atrasos e 19446 pontos de adiantamentos. No cenário  $B_z \times B_{z,rMHD}$  obteve-se uma distância Euclidiana 35,6, DTW 0,2 e RMSE 0,98. Já no cenário  $B_z \times B_{z,hMHD}$  obteve-se distância Euclidiana 24,7, DTW 0,13 e RMSE 0,68. Observa-se que os valores para as métricas obtidos estabelecem diferenciações entre os desempenhos dos modelos. De certa forma, o modelo Hall (hMHD) apresenta valores em maior proximidade aos valores medidos do que os do modelo puramente resistivo. A Tabela 6.1 apresenta os resultados compilados dos cenários analisados para a similaridades entre os dados observacionais e dos modelos rMHD e hMHD referentes da componente do campo magnético  $B_z$ . Para da completeza ao estudo de variáveis físicas, uma tabela completa considerando outras componentes magnéticas está apresentada no Apêndice B.

Tabela 6.1 - Similaridades entre os dados observacionais e dos modelos rMHD e hMHD referentes à componente do campo magnético  $B_z$ .

Séries	Caminho W (área)		Distância		RMSE
	Atrasos	Adiantamentos	Euclidiana	DTW	
$[B_z, B_{z,rMHD}]$	122.445	223.494	35,6	0,20	0,98
$[B_z, B_{z,hMHD}]$	44.202	190.446	24,7	0,13	0,68

Fonte: Produção do Autor.

No caso dessas simulações, a interpretação dos deslocamentos temporais diz respeito às características dos processos físicos simulados não estarem adequadamente retratados em confronto com as avaliações físicas (medidas) coletadas pelo satélite. O comportamento comparado entre os modelos revelam que embora o início traga similaridade, à medida que o tempo passa as respostas tendem a apresentar características próprias. O método, como está, não serve para escolher uma melhor resposta;

no entanto, em complementação às análises por técnicas *wavelets*, identifica necessidades de aprimoramentos das modelagens numérico-computacionais envolvidas. Em condições adequadas de modelagem, os sinais deveriam apresentar um mínimo ou nenhum deslocamento temporal.

## **6.2 Aplicação da metodologia a dados classificadores de distúrbios geomagnéticos**

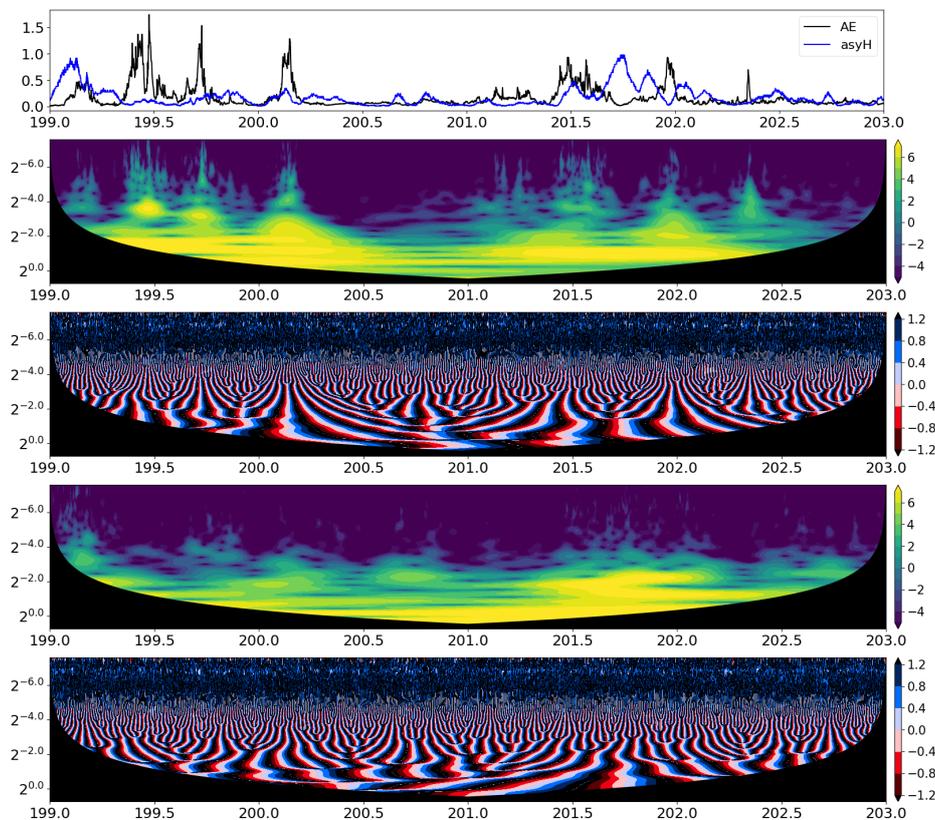
Nesta seção, analisam-se dados classificadores das respostas de distúrbio geomagnético avaliados por equipamentos instalados na superfície terrestre. Esses distúrbios, que se caracterizam principalmente como tempestades e subtempestades geomagnéticas, decorrem de processos físicos de acoplamento eletrodinâmico entre o plasma magnetizado interplanetário e o ambiente terrestre. De forma a implementar uma metodologia para uma melhor interpretação de inter-relação das regiões de altas latitudes ( $> 50$  graus) e baixas latitudes ( $< 30$  graus), escolheram-se para análise os índices geomagnéticos AE, resposta auroral, e ASY-H, resposta quase equatorial.

Ainda referente a esses dados em estudos, selecionaram-se dois períodos típicos caracterizadores de acoplamentos eletrodinâmicos. Um intervalo referente a período geomagneticamente calmo, designado abreviadamente período calmo, compreendendo os dias 18 a 21 de julho (DOY 199 a 202) de 2006. O outro intervalo referente a um período perturbado, decorrente de um evento CIR no meio interplanetário, compreendendo os dias 12 a 15 de fevereiro (DOY 43 a 46) de 2004. Por causa das amplitudes (dadas em nanoteslas) muito diferenciadas dos índices, os dados foram normalizados e multiplicados aqui convenientemente por um fator 1000. Este procedimento numérico não afeta em nada os comportamentos físicos implícitos nos índices. Desta forma torna-se possível avaliar as relações dos sinais e suas variabilidades sem a interferência da diferença conceitual de amplitude dos índices geomagnéticos. Busca-se identificar as principais características desses sinais e suas relações, considerando, para completeza da interpretação metodológica, períodos geomagneticamente distintos.

### **6.2.1 Análise dos dados para período geomagneticamente calmo**

Analisam-se os índices geomagnéticos AE e ASY-H no período geomagneticamente calmo. As características tempo-escala dos dados estão apresentadas na Figura 6.10. Os tratamentos metodológicos e as equivalentes leituras nesta parte são idênticos aos descritos na seção anterior. Então, de forma mais objetiva, atenta-se aqui e discutem-se as partes essenciais.

Figura 6.10 - Escalogramas e fases do período calmo dos Dados dos índices AE e ASY-H.



Fonte: Produção do Autor.

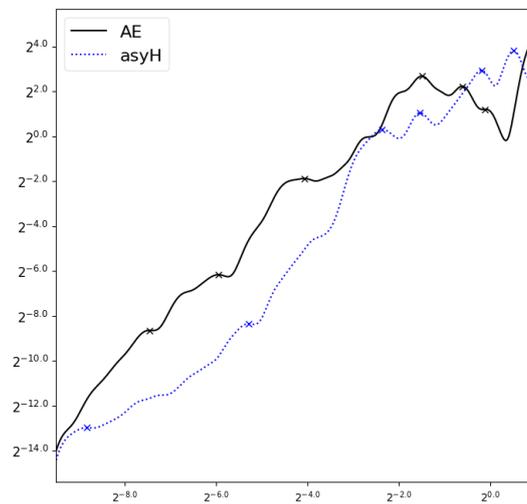
A figura traz, no painel ao topo, os índices classificadores. Logo abaixo estão o escalograma e o gráfico de fases do AE e mais abaixo, esses resultados para o ASY-H.

Nos classificadores, notam-se os intervalos em que perturbações, embora fracas comparadas a períodos verdadeiramente perturbados, manifestam-se. Alguma relação há, em termos de intensidade e intervalo temporal, entre essas variações nos índices; como também há distinções aparentes. A metodologia visa destacar as características gerais e particulares desses sinais. No escalograma do índice AE, observam-se estruturas de alta energia do sinal localizadas entre as escalas  $2^{-2}$  (6 horas) e  $2^0$  (1 dia) presentes e mais contínuas praticamente ao longo de todo o intervalo. No ASY-H também a essa caracterização. Contudo, em torno do período maior ( $\approx 1$  dia, ou frequência menor), a contribuição em energia mostra-se maior em ASY-H do que no AE. Surtos de variação magnética são notados, particularmente para o AE, em períodos menores (frequências mais altas). Assim, nota-se que durante as variações

que dão as maiores amplitudes para o sinal localizadas entre os intervalos dos dias 199 a 200,2 e 201 a 202,4, surgem no escalograma estruturas de alta energia entre as escalas que representam 0,3 a 5 horas. Constatam-se objetivamente inter-relações entre esses classificadores, que devem estar relacionados a processos de larga escala espacial. A análise de fase, embora no geral se assemelhe para os classificadores, indica comportamento mais assemelhando nas médias frequências (períodos baixos) e que vai se distinguindo nas baixa frequências (períodos altos).

A Figura 6.11 apresenta o espectro global *wavelet* para AE e ASY-H, no período calmo.

Figura 6.11 - Espectro global *wavelet* para AE e ASY-H, no período calmo.



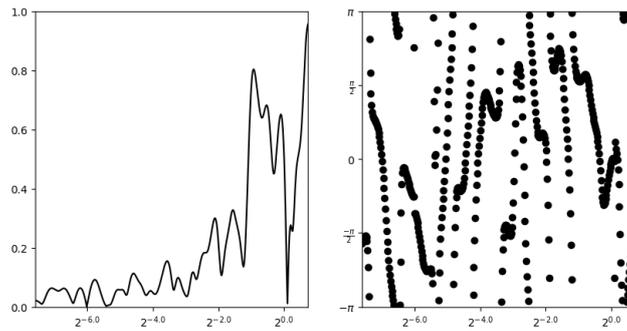
Fonte: Produção do Autor.

O gráfico mostra os picos aparentes, identificando intervalos de frequências predominantes. Nos períodos mais curtos,  $< 6$  h (frequências mais longas), há uma ocorrência de picos algo próximas nos dois sinais, embora a energia relativa em ASY-H seja menor que em AE. Enquanto, nos períodos mais longos que 6 h, ocorram picos nos dois classificadores sob padrões mais complexos. Para ASY-H, há um crescimento monotônico da energia dos picos; já para AE há uma flutuação. É possível analisar nos espectros globais que os sinais apresentam poucas escalas com picos de energia em comum. Há similaridades nas curvas dos espectros, esse comportamento é visível nos ciclos de maiores escalas. Interpreta-se que esse diagnóstico não apenas confirma frequências mais comuns no processo de acoplamento, como

também distinções na morfologia dos sinais estabelecidos.

Continuando a análise dos dados, a Figura 6.12 traz as correlações por escala, em que o módulo é apresentado no painel à esquerda e a fase à direita.

Figura 6.12 - Período calmo AE e ASY-H: correlação por escalas e fase.

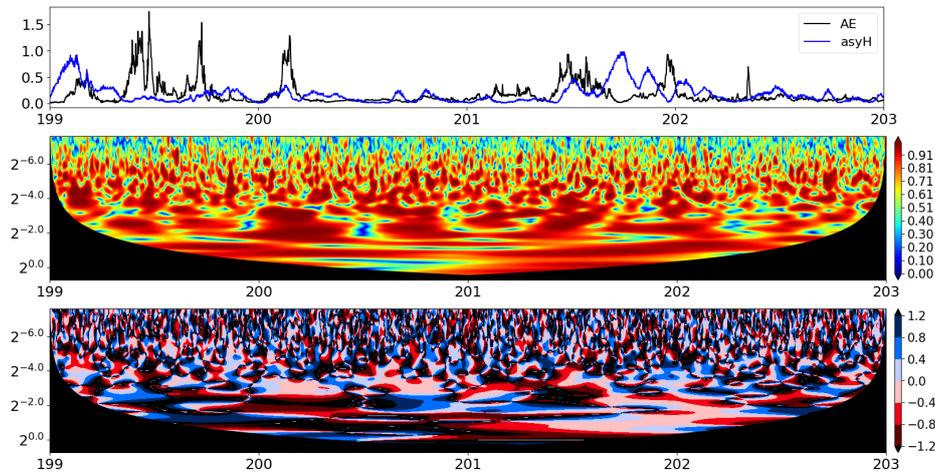


Fonte: Produção do Autor.

Embora manifestem-se correlações ao longo do intervalo de períodos, algum significado pode ser considerado para períodos longos a partir de quase 6 horas. Alguns intervalos de períodos (frequências) se destacam,  $R^2 \geq 50\%$ , em 6h, 9h, 12h e 2 dias, cuja análise de fase denotam estar praticamente em fase.

Na Figura 6.13 está apresentada a coerência *wavelet* e diferença de fase dos índices AE e ASY-H, no período calmo, apenas nas escalas próximas as maiores correlações.

Figura 6.13 - Coerência *wavelet* e diferença de fase para AE e ASY-H no período calmo.

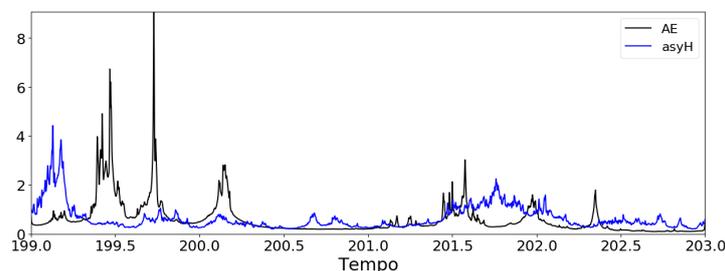


Fonte: Produção do Autor.

Na análise de coerência é possível observar que ocorre, predominantemente, coerência nas escalas próximas as 6 horas a 2 dias ao longo de todo o intervalo. Na diferença de fases, fica visível a predominância de fase, conquanto haja nas várias escalas intervalos temporais com adiantamento ou retardo de um sinal com respeito ao outro. Esses deslocamentos temporais caracterizam uma complexidade dos mecanismos de injeção de energia durante condição de calma magnética.

Procede-se agora com a análise da entropia dos sinais, apresentada na Figura 6.14.

Figura 6.14 - Entropia *wavelet* dos índices AE e ASY-H no período calmo.



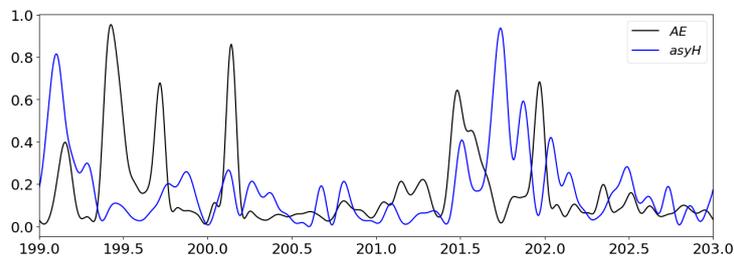
Fonte: Produção do Autor.

Para ambos os sinais, os picos de entropia ocorrem nos momentos de maior amplitude do sinal. De detalhes explicitados na Figura 6.10, verifica-se que as frequências mai-

ores (períodos menores) estão relacionados particularmente a esses incrementos de energia no interior da magnetosfera. Em contraposição, em situação que mereceria exame físico, o intervalo temporal posterior ao DOY 201, manifestam-se ocorrências com entropias similares, não muito pronunciadas e desenvolvendo-se com uma duração mais dilatada.

Por último, a Figura 6.15 apresenta a reconstrução dos dados, considerando uma filtragem que destacassem as escalas predominantes nos sinais, conforme as análises anteriores.

Figura 6.15 - Sinais reconstruídos por filtragem *wavelet* para os índices AE e ASY-H no período calmo.



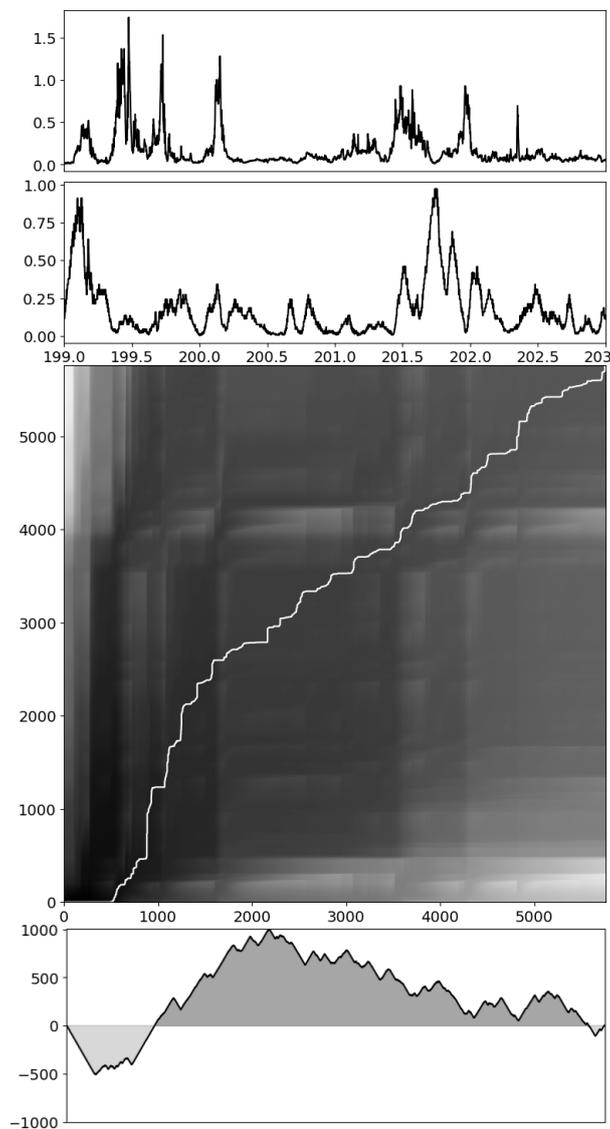
Fonte: Produção do Autor.

Os comportamentos explicitados nesta representação dos classificadores AE e ASY-H capturam o papel dos ciclos diurnos e semi-diurnos em predomínio sobre os efeitos externos. Contudo a forma de contribuição das energias distribuídas entre o sistema equatorial (ASY-H) e auroral (AE) não mostram plena concordância de comportamento.

#### 6.2.1.1 Análise das métricas de similaridade entre dados dos índices geomagnético no período geomagneticamente calmo

Com a aplicação da DTW, avaliam-se os deslocamentos temporais relacionados aos índices geomagnéticos AE e ASY-H, apresentados na Figura 6.16.

Figura 6.16 - Matriz de Distância e Caminho  $\mathfrak{W}$  para AE e ASY-H, no período calmo.



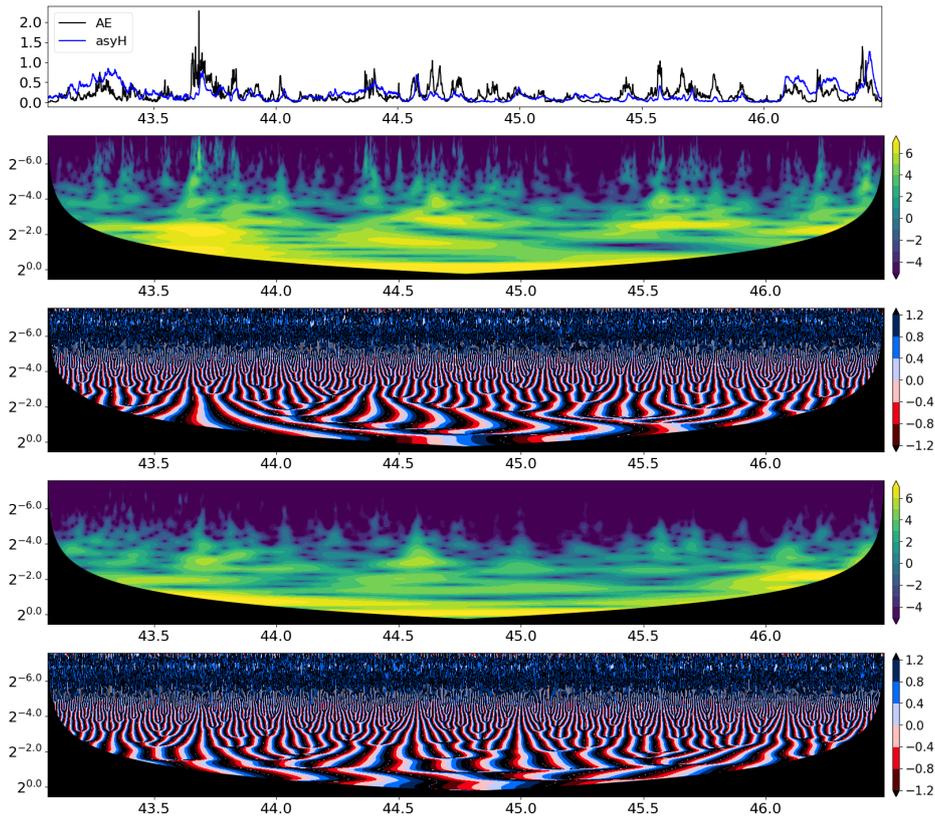
Fonte: Produção do Autor.

Para os classificadores para a condição geomagneticamente calma, obtiveram-se áreas de 512.004 pontos de atraso de AE para ASY-H e 3.620.677 pontos de adiantamento. Nesta análise, onde os processos por trás de ambos os sinais são mais estocásticos, o índice AE apresentou adiantamentos nos picos de variação, esse comportamento fica visível nos sinais filtrados nas principais escalas, apresentado em 6.15. Calcularam-se as seguintes métricas de similaridade: o valor de  $\mathcal{D}$  é 20,9, para  $\mathcal{Q}$  é 307,8 e para RMSE é 0,28, que estão apresentadas na Tabela 6.2.

### 6.2.2 Análise dos dados em período geomagneticamente perturbado

Analisa-se os índices geomagnéticos AE e ASY-H no período geomagneticamente perturbado decorrente de evento CIR no meio interplanetário. As características tempo-escala dos dados estão apresentadas na Figura 6.17.

Figura 6.17 - Escalograma e fases dos índices AE e ASY-H em período perturbado.



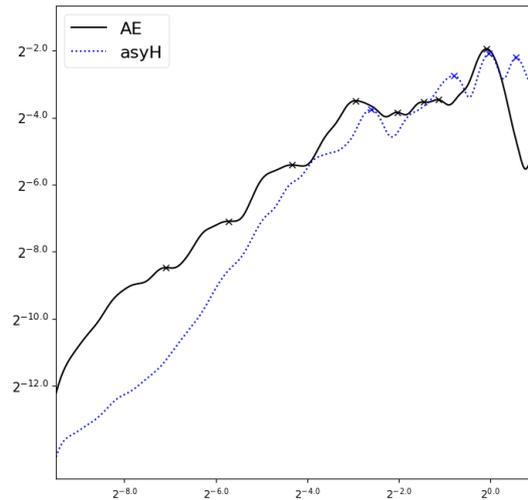
Fonte: Produção do Autor.

No topo estão os índices classificadores, a seguir o escalograma e o gráfico de fases do AE e abaixo, os resultados para o ASY-H. Diferentemente do período calmo, o período perturbado apresenta estruturas de energias em escalas de variações mais rápidas em ambos os índices durante todos os dias analisados. Essas características de perturbação nos índices são consequência de maior efetividade na transferência de energia entre as partes da magnetosfera, pelo papel preponderante da componente  $B_z$  ao longo do fenômeno CIR. Nota-se nos escalogramas maior ocorrência de contribuições nos sinais das escalas de baixo período (maior frequência). Por outro lado, ao longo da maior parte das escalas há também maior contribuição de energia.

As fases revelam uma similaridade; embora não sejam realmente idênticas.

A Figura 6.18 apresenta o espectro global *wavelet* para AE e ASY-H, no período perturbado.

Figura 6.18 - Espectro global *wavelet* para AE e ASY-H, no período perturbado.

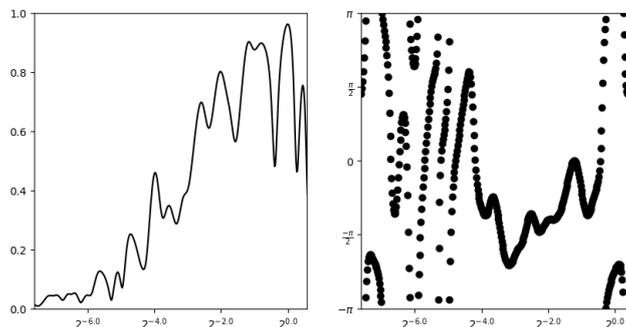


Fonte: Produção do Autor.

Nesta figura é possível observar que, em relação ao período calmo, o período perturbado possui uma maior contribuição de energia nos períodos mais longos (frequências mais altas). O platô de contribuição de energia para os dois sinais são mais coincidentes no intervalo, período  $\geq 3$  h, com alguma variação.

Analisa-se as correlações por escala (módulo à esquerda e fase à direita) dos dados na Figura 6.19.

Figura 6.19 - Correlação por escalas e fase para AE e ASY-H no período perturbado.

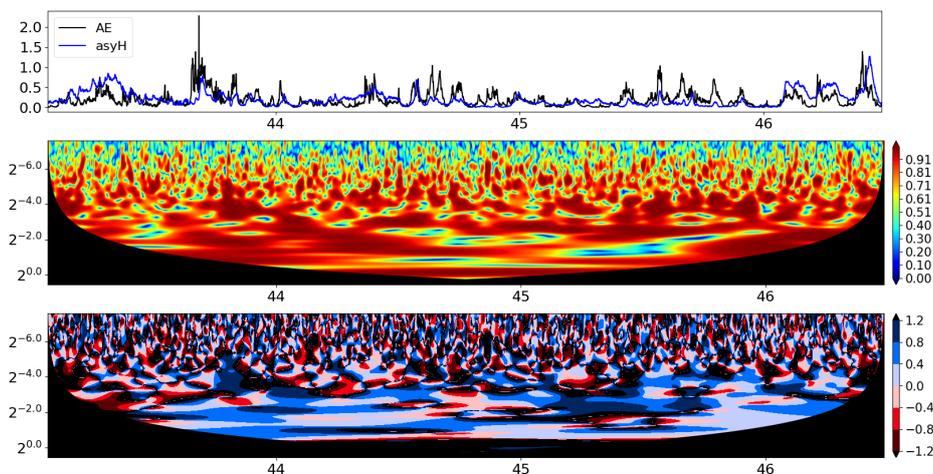


Fonte: Produção do Autor.

Há maiores correlações em mais escalas com tendência de comportamento em fase nessas escalas. Visto que há um evento físico gerador do fenômeno CIR influenciando ambos os dados esse comportamento menos estocástico é esperado. Nas escalas de mais baixas energias, as de variações rápidas, o sinal continua apresentando um comportamento menos uniforme, podendo que concluir que o fenômeno CIR impacta apenas determinadas escalas.

Na Figura 6.20 está apresentada a coerência *wavelet* e diferença de fase dos índices apenas nas escalas próximas as maiores correlações.

Figura 6.20 - Coerência *wavelet* e diferença de fase para AE e ASY-H, no período perturbado.

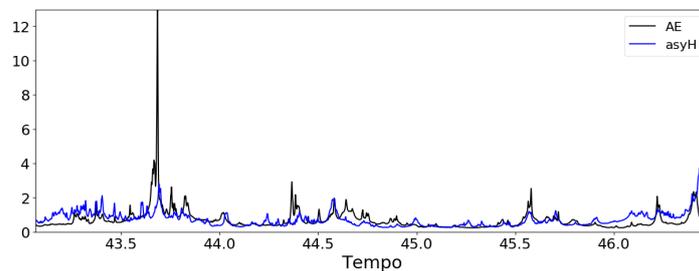


Fonte: Produção do Autor.

A coerência *wavelet* apresenta grandes estruturas entre as escalas 1,5 h e 2 dias, comportamento observado nas correlações por escala. Destaca-se principalmente as estruturas entre os dias 43 a 45 e depois próximas ao dia 46, onde estão as maiores relações. Na análise de diferença de fase observa-se que próximo as estruturas citadas acima é visível a condução do sinal do índice AE, desta forma conclui-se que AE está adiantado em relação ao sinal de ASY-H.

O resultado da análise da entropia dos sinais é mostrada na Figura 6.21.

Figura 6.21 - Entropia *wavelet* para AE e ASY-H, período perturbado.

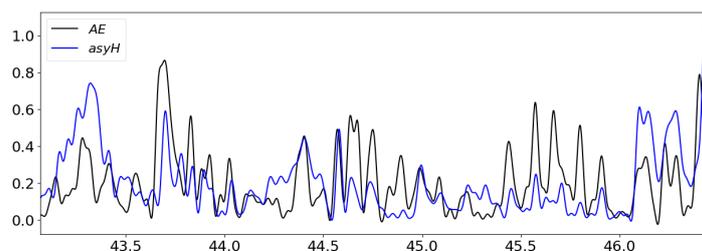


Fonte: Produção do Autor.

Na entropia *wavelet* de ambos os sinais, é possível observar que há picos em comum, com uma leve tendência de adiantamento dos picos na série de AE, nesta série é possível observar um pico que destoa do resto da série temporal, dada a característica de AE de atingir maiores amplitudes em relação ao ASY-H.

Na Figura 6.22 estão apresentados as reconstruções dos sinais, com base nos períodos destacados pela análise *wavelet* com respeito às maiores correlações e coerência.

Figura 6.22 - Sinais reconstruídos pela técnica *wavelet* para AE e ASY-H.



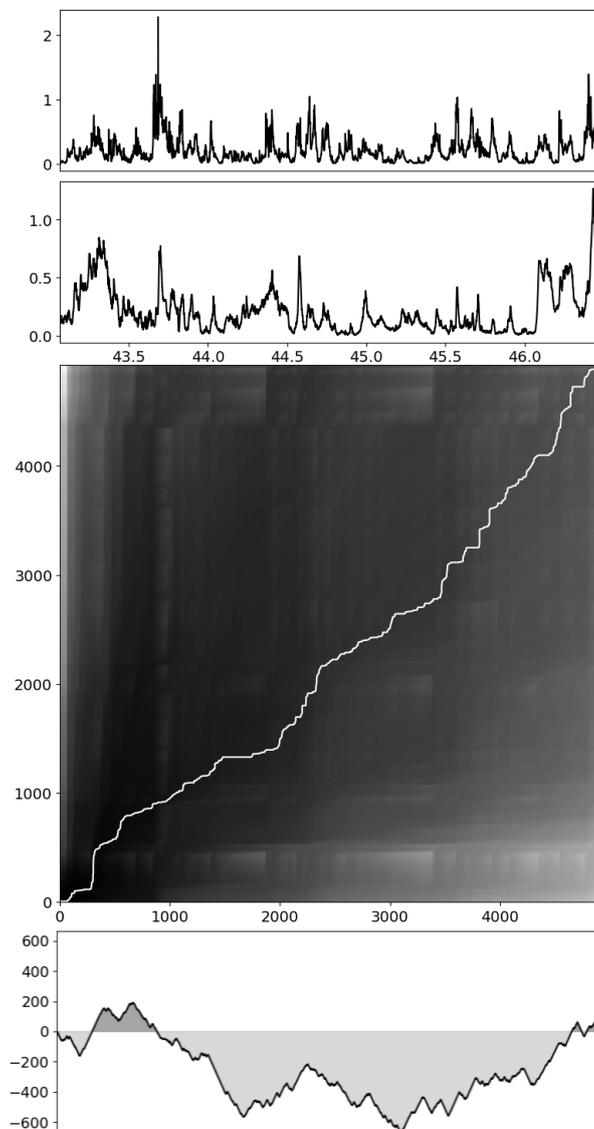
Fonte: Produção do Autor.

Analisando os sinais reconstruídos é possível observar uma maior quantidade de picos e vales localizados nas mesmas posições temporais, confirmando o comportamento menos estocástico.

### 6.2.2.1 Análise de métricas de similaridade entre dados dos índices geomagnéticos no período geomagneticamente perturbado

Utilizando a técnica DTW, analisam-se os deslocamentos temporais, que são mostrados na Figura 6.23.

Figura 6.23 - Matriz de Distância e Caminho  $\mathfrak{W}$  para AE e ASY-H, para período perturbado.



Fonte: Produção do Autor.

Na condição de perturbação geomagnética, obtiveram-se as áreas de 2.221.510 pontos de atraso e 113.337 pontos de adiantamento de AE em relação a ASY-H. Nesta situação, em que há um fenômeno físico mais intenso que influencia ambos os sinais, observa-se menor número de deslocamentos entre os dois sinais. Constatou-se, no entanto, que o sinal do índice AE apresenta defasagem nos picos dos fenômenos.

Na condição de perturbação, obtiveram-se as seguintes métricas de similaridade: o valor de  $\mathcal{D}$  é 15, 24, para  $\mathcal{Q}$  é 283, 4 e para RMSE é 0, 22. A Tabela 6.2 apresenta esses valores para a condição geomagneticamente perturbada, como também os valores para a condição calma.

Tabela 6.2 - Similaridades entre os dos índices geomagnéticos AE e ASY-H no período geomagneticamente perturbado.

Séries	Caminho W (área)		Distância		RMSE
	Atrasos	Adiantamentos	Euclidiana	DTW	
Período Calmo	512.004	3.620.677	20, 91	307, 8	0, 28
Período Perturbado	2.221.510	113.337	15, 24	283, 4	0, 22

Fonte: Produção do Autor.

Resumindo em um comentário mais geral, por meio dos sinais reconstruídos pela técnica *wavelet*, apresentadas nas Figuras 6.15 e 6.22, observa-se que nas escalas do caso perturbado, ao contrário do caso calmo, onde há maiores correlações as séries temporais possuem bom encaixe temporal no mapeamento dos picos e vales.

Observa-se também que o período perturbado, em diferença com o calmo, apresenta valores menores nas métricas de similaridade. Tal aspecto é condiz com a expectativa, visto que há um fenômeno físico que contribui com determinada faixa de escala de ambos os sinais, tornando os processos entre ambos menos estocásticos.

Por último, da comparação entre os casos calmo e perturbado, destaca-se principalmente as diferenças nos atrasos e adiantamentos. No período perturbado, onde apresenta-se o maior pico na série temporal AE e em suas proximidades, é onde há o menor número de deslocamentos temporais na DTW.

A metodologia implementada neste trabalho sinaliza um potencial uso para muitas outras abordagens no contexto das Ciências espaciais, estendendo-se mesmo para

propósitos em outras áreas.



## 7 CONCLUSÃO

Definiu-se como principal objetivo deste trabalho a criação de uma metodologia envolvendo elementos multirresolução e a implementação de recursos numérico-computacionais para análise de dados espaciais.

Foram utilizadas ferramentas de análise *wavelet* para caracterização de fenômenos e identificação de relações tempo-espaco entre dois sinais. Também foram utilizadas métricas e ferramentas de análise de similaridade física para quantificação global.

Neste trabalho, realizaram-se as, por meio da utilização de diversos dados sintéticos, as validações das ferramentas selecionadas e preparadas para constituir a metodologia integrada de análises. Como casos de aplicações, procederam-se, em uma primeira parte, as análises de dados de simulações de modelos MHD resistivo e Hall em conjunto com dados experimentais coletados pela sonda Cassini, da NASA, nas proximidades do satélite natural Titã do planeta Saturno, o que permitiu avaliar a representação por modelo das condições eletrodinâmicas de um astro. Em uma segunda parte, a mesma metodologia foi utilizada para análises de medidas referentes ao acoplamento eletrodinâmico do plasma do vento solar e da magnetosfera terrestre, para o qual se empregaram quantificações das perturbações geomagnéticas registradas na superfície.

A metodologia desenvolvida com este trabalho consiste das seguintes características e etapas de aplicação:

- a) identificação dos principais aspectos tempo-escala dos sinais analisados, utilizando o escalograma, espectro global e fase;
- b) análise das relações tempo-escala, utilizando a correlação por escala, coerência *wavelet* e diferença de fase;
- c) decomposição dos sinais e reconstrução em escalas equiparáveis, utilizando a abordagem MODWT;
- d) quantificação das similaridades, utilizando a distância Euclidiana, RMSE e DTW; e
- e) análise da quantificação de distorções temporais dinâmicas entre séries, utilizando a técnica DTW.

Nos resultados preliminares das ferramentas multiescala ficaram claros os potenciais analíticos das ferramentas *wavelet*, em que foi possível identificar as estruturas de sinais sintéticos e, principalmente, de sinais já estudados na literatura.

Caracterizaram-se as principais diferenças entre a DWT e a MODWT, em que foi possível identificar as vantagens na utilização da MODWT para obter uma reconstrução mais rica em detalhes locais de sinais.

Foi possível observar pelos testes de sensibilidade que a ferramenta DTW é robusta para estudos de diferentes tipos de sinais, em que se avaliaram as diferenças em sinais de frequências, fase e localização de eventos distintos.

No estudo dos dados de modelagens MHD considerando os efeitos dos termos resistivo (rMHD) e Hall (hMHD) foi possível observar, como esperado, que os dados simulados não possuem a mesma riqueza de detalhamento tempo-escala dos sinais experimentais. As componentes de maior escala do sinal são reproduzidas em ambos os casos. Adicionalmente, o modelo hMHD apresentou um melhor aspecto tempo-escala com tendência mais próxima ao das medições.

Por último, a análise dos índices geomagnéticos destacaram, por meio da metodologia, a distinção entre processos físicos com estruturas de injeção de energias mais organizadas, como em tempestades geomagnéticas, em relação a processos de fraca estrutura organizada, como períodos geomagneticamente calmos.

### **Perspectivas de Trabalhos Futuros**

Como sugestão para futuros trabalhos, destaca-se a implementação de um serviço web para análise de dados, com facilitação de *input*, robustez de processo e consistência de *output*.

Sugere-se também uma ampliação deste estudo para outras situações de modelagens, em que se confrontem dados observacionais com dados de modelo.

Outra possibilidade é a implementação de recursos computacionalmente inteligentes para a identificação de similaridades em coeficientes *wavelet* e reconstrução de sinais equiparáveis entre si sem interação humana.

## REFERÊNCIAS BIBLIOGRÁFICAS

AGUIAR, C.; SOARES, M. J. The continuous wavelet transform: moving beyond uni-and bivariate analysis. **Journal of Economic Surveys**, v. 28, n. 2, p. 344–375, 2014. 33, 99

ANTOINE, J. P.; MURENZI, R.; VANDERGHEYNST, P.; ALI, S. T. **Two-dimensional wavelets and their relatives**. [S.l.]: Cambridge University Press, 2008. 19, 20, 21

BARNSTON, A. G. Correspondence among the correlation, rmse, and heidke forecast verification measures; refinement of the heidke score. **Weather and Forecasting**, v. 7, n. 4, p. 699–709, 1992. 49

BERNDT, D. J.; CLIFFORD, J. Using dynamic time warping to find patterns in time series. In: AIAA WORKSHOP ON KNOWLEDGE DISCOVERY IN DATABASES. **Proceedings...** [S.l.]: AIAA, 1994. v. 10, n. 16, p. 359–370. 50

BITTENCOURT, J. A. **Fundamentals of plasma physics**. [S.l.]: Springer Science & Business Media, 2013. 10

CAMPBELL, W. H. **Introduction to geomagnetic fields**. [S.l.]: Cambridge University Press, 2003. 15

CROOKER, N. U.; SISCOE, G. L. A study of the geomagnetic disturbance field asymmetry. **Radio Science**, v. 6, n. 4, p. 495–501, 1971. 16

DAUBECHIES, I. The wavelet transform, time-frequency localization and signal analysis. **IEEE Transactions on Information Theory**, v. 36, n. 5, p. 961–1005, 1990. 46

\_\_\_\_\_. **Ten lectures on wavelets**. [S.l.]: CBMS-NSF Regional Conference Series in Applied Mathematics, 1992. 21

DAVIDSON, P. A. **Introduction to magnetohydrodynamics**. [S.l.]: Cambridge University Press, 2001. 10

DOMINGUES, M. O. **Análise wavelet na simulação numérica de equações diferenciais parciais com adaptabilidade espacial**. PhD Thesis (PhD) — Universidade de Campinas, Campinas, 2001. 40, 42

DOMINGUES, M. O.; MENDES, O.; COSTA, A. M. D. On wavelet techniques in atmospheric sciences. **Advances in Space Research**, v. 35, n. 5, p. 831–842, 2005. [17](#), [18](#)

DOMINGUES, M. O.; MENDES, O.; KAIBARA, M. K.; MENCONI, V. E.; BERNARDES, E. Explorando a transformada wavelet contínua. **Revista Brasileira de Ensino de Física**, v. 38, n. 3, 2016. [36](#)

ESA. Solar wind. 2018. [2](#)

FALOUTSOS, C.; RANGANATHAN, M.; MANOLOPOULOS, Y. **Fast subsequence matching in time-series databases**. [S.l.]: ACM, 1994. [49](#)

FARGE, M. Wavelet transforms and their applications to turbulence. **Annual Review of Fluid Mechanics**, v. 24, n. 1, p. 395–458, 1992. [19](#), [21](#), [23](#)

FRICK, P.; BECK, R.; BERKHUIJSEN, E. M.; PATRICKEYEV, I. T. Scaling and correlation analysis of galactic images. **Monthly Notices of the Royal Astronomical Society**, v. 327, n. 4, p. 1145–1157, 2001. [33](#)

FRIIS-CHRISTENSEN, E.; LASSEN, K. Length of the solar cycle: an indicator of solar activity closely associated with climate. **Science**, v. 254, n. 5032, p. 698–700, 1991. [24](#)

FU, T. A review on time series data mining. **Engineering Applications of Artificial Intelligence**, v. 24, n. 1, p. 164–181, 2011. [49](#)

GOEDBLOED, J. P.; POEDTS, S. **Principles of magnetohydrodynamics: with applications to laboratory and astrophysical plasmas**. [S.l.]: Cambridge University Press, 2004. [10](#)

GONZALEZ, W. D.; JOSELYN, J. A.; KAMIDE, Y.; KROEHL, H. W.; ROSTOKER, G.; TSURUTANI, B. T.; VASYLIUNAS, V. M. What is a geomagnetic storm? **Journal of Geophysical Research: Space Physics**, v. 99, n. A4, p. 5771–5792, 1994. [16](#)

GROSSMANN, A.; MORLET, J. Decomposition of hardy functions into square integrable wavelets of constant shape. **SIAM Journal on Mathematical Analysis**, v. 15, n. 4, p. 723–736, 1984. [18](#), [20](#)

HUDGINS, L.; FRIEHE, C. A.; MAYER, M. E. Wavelet transforms and atmospheric turbulence. **Physical Review Letters**, v. 71, n. 20, p. 3279, 1993. [32](#)

- IYEMORI, T. Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations. **Journal of Geomagnetism and Geoelectricity**, v. 42, n. 11, p. 1249–1265, 1990. 16
- JACH, A.; KOKOSZKA, P.; SOJKA, J.; ZHU, L. Wavelet-based index of magnetic storm activity. **Journal of Geophysical Research: Space Physics**, v. 111, n. A9, 2006. 45
- KABATA, W.; VITORELLO, I. Configurações alternativas para magnetômetros "fluxgate" com núcleo amorfo. **Revista Brasileira de Geofísica**, v. 25, n. 2, p. 159–169, 2007. 15
- KAMIDE, Y.; CHIAN, A. **Handbook of the solar-terrestrial environment**. [S.l.: s.n.], 2007. 1
- KEOGH, E.; CHAKRABARTI, K.; PAZZANI, M.; MEHROTRA, S. Dimensionality reduction for fast similarity search in large time series databases. **Knowledge and Information Systems**, v. 3, n. 3, p. 263–286, 2001. 51
- KEOGH, E.; RATANAMAHATANA, C. A. Exact indexing of dynamic time warping. **Knowledge and Information Systems**, v. 7, n. 3, p. 358–386, 2005. 50, 52
- KIRCHOFF, V. W. J. H. **Introdução a geofísica espacial**. [S.l.: s.n.], 1991. 8
- KIVELSON, M. G.; RUSSELL, C. T. **Introduction to space physics**. [S.l.]: Cambridge University Press, 1995. 2, 7, 13, 15
- LABAT, D. Recent advances in wavelet analyses: Part 1. a review of concepts. **Journal of Hydrology**, v. 314, n. 1-4, p. 275–288, 2005. 32
- LAU, K. M.; WENG, H. Climate signal detection using wavelet transform: how to make a time series sing. **Bulletin of the American Meteorological Society**, v. 76, n. 12, p. 2391–2402, 1995. 17, 18
- LIN, J.; KHADE, R.; LI, Y. Rotation-invariant similarity in time series using bag-of-patterns representation. **Journal of Intelligent Information Systems**, v. 39, n. 2, p. 287–315, 2012. 50, 51
- LORENZO, M. S.; DOMINGUES, M. O.; MECÍAS, A. L.; MENCONI, V. E.; MENDES, O. On the use of space-environmental satellite data for global magnetohydrodynamic simulations. **Brazilian Journal of Physics**, v. 46, n. 6, p. 703–713, 2016. 2

MA, Y.; NAGY, A. F.; TOTH, G.; CRAVENS, T. E.; RUSSELL, C. T.; GOMBOSI, T. I.; WAHLUND, J.; CRARY, F. J.; COATES, A. J.; BERTUCCI, C. L.; NEUBAUER, F. M. 3D global multi-species Hall-MHD simulation of the Cassini T9 flyby. **Geophysical Research Letter**, v. 34, p. L24S10:1–6, 2007. [12](#), [13](#), [64](#)

MALLAT, S. G. A theory for multiresolution signal decomposition: the wavelet representation. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v. 11, n. 7, p. 674–693, 1989. [38](#)

MEJNERTSEN, L.; EASTWOOD, J. P.; HIETALA, H.; SCHWARTZ, S. J.; CHITTENDEN, J. P. Global mhd simulations of the earth's bow shock shape and motion under variable solar wind conditions. **Journal of Geophysical Research: Space Physics**, v. 123, n. 1, p. 259–271, 2018. [3](#)

MENDES, O. **A origem interplanetária eo desenvolvimento da fase principal das tempestades geomagnéticas moderadas (1978-1979)**. Tese (Doutorado em Geofísica Espacial) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1992. [1](#), [7](#), [13](#)

MENDES, O.; COSTA, A. M.; DOMINGUES, M. O. Introduction to planetary electrodynamics: a view of electric fields, currents and related magnetic fields. **Advances in Space Research**, v. 35, n. 5, p. 812–828, 2005. [7](#), [8](#), [13](#)

MENDES, O.; DOMINGUES, M. O.; COSTA, A. M.; GONZALEZ, A. L. C. D. Wavelet analysis applied to magnetograms: singularity detections related to geomagnetic storms. **Journal of Atmospheric and Solar-Terrestrial Physics**, v. 67, n. 17-18, p. 1827–1836, 2005b. [2](#)

MOLDWIN, M. **An introduction to space weather**. [S.l.]: Cambridge University, 2008. [1](#), [2](#)

MORLET, J. Sampling theory and wave propagation. **Issues in Acoustic Signal—image Processing and Recognition**, p. 233–261, 1983. [20](#)

MULLER, M. Dynamic time warping. **Information Retrieval for Music and Motion**, p. 69–84, 2007. [50](#)

NASON, G. **Wavelet methods in statistics with R**. [S.l.]: Springer Science & Business Media, 2008. [44](#)

- NOSE, M.; T., I.; ODAGI, Y.; TAKEDA, M.; TOH, H.; KOYAMA, Y.; TAKEUCHI, N. **Data analysis center for geomagnetism and space magnetism**. [S.l.]: Graduate School of Science, Kyoto University, Kyoto, Japan, 2011. 15
- OGINO, T.; WALKER, R. J.; ASHOUR-ABDALLA, M. An mhd simulation of the interaction of the solar wind with the outflowing plasma from a comet. **Geophysical Research Letters**, v. 13, n. 9, p. 929–932, 1986. 1, 11, 12
- PERCIVAL, D. B.; WALDEN, A. T. **Wavelet methods for time series analysis**. [S.l.]: Cambridge University Press, 2000. 40, 41, 42, 44
- QUIROGA, R. Q.; ROSSO, O. A.; BAŞAR, E.; SCHÜRMAN, M. Wavelet entropy in event-related potentials: a new method shows ordering of eeg oscillations. **Biological Cybernetics**, v. 84, n. 4, p. 291–299, 2001. 30
- RAFIEL, D.; MENDELZON, A. Similarity-based queries for time series data. **ACM SIGMOD Record**, v. 26, n. 2, p. 13–25, 1997. 49
- ROBERTS, P. H. **An introduction to magnetohydrodynamics**. [S.l.]: Longmans London, 1967. 10
- RUSSELL, C. T.; MCPHERRON, R. L. Semiannual variation of geomagnetic activity. **Journal of Geophysical Research**, v. 78, n. 1, p. 92–108, 1973. 24
- SAKOE, H. Dynamic programming algorithm optimization for spoken word recognition. **IEEE Transactions on Acoustic Speech and Signal Processing**, v. 26, n. 1, p. 43–49, 1978. 50
- SCHUNK, R.; NAGY, A. **Ionospheres: physics, plasma physics, and chemistry**. [S.l.]: Cambridge University Press, 2009. 8
- SELLO, S. Wavelet entropy and the multi-peaked structure of solar cycle maximum. **New Astronomy**, v. 8, n. 2, p. 105–117, 2003. 30
- SEO, R. T.; DOMINGUES, M. O.; MENDES, O. Análise de parâmetros do meio interplanetários utilizando a transformada wavelet cruzada. **Proceeding Series of the Brazilian Society of Computational and Applied Mathematics**, v. 6, n. 2, 2018. 34
- SHUE, J.-H.; SONG, P.; RUSSELL, C. T.; STEINBERG, J. T.; CHAO, J. K.; ZASTENKER, G.; VAISBERG, O. L.; KOKUBUN, S.; SINGER, H. J.; DETMAN,

T. R. Magnetopause location under extreme solar wind conditions. **Journal of Geophysical Research: Space Physics**, v. 103, n. A8, p. 17691–17700, 1998. 3

TORRENCE, C.; COMPO, G. P. A practical guide to wavelet analysis. **Bulletin of the American Meteorological society**, v. 79, n. 1, p. 61–78, 1998. 18

TORRENCE, C.; WEBSTER, P. J. Interdecadal changes in the enso–monsoon system. **Journal of Climate**, v. 12, n. 8, p. 2679–2690, 1999. 32

YAMASHITA, C. S. **Efeito das tempestades magnéticas intensas na ionosfera de baixa latitude**. Dissertação (Mestrado em Geofísica Espacial) — Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 1999. 15

## Apêndice A: Estudos sobre os efeitos de suavização da coerência *wavelet*

A seguir, para analisar as principais diferenças no processo de suavização da coerência *wavelet*, utilizam-se dois sinais que possuem dois ciclos em comum com uma defasagem temporal. As séries simulam dados mensais em um período de 50 anos adicionados de um ruído  $\epsilon$  aleatório, apresentadas nas Equação A.1 e Equação A.2.

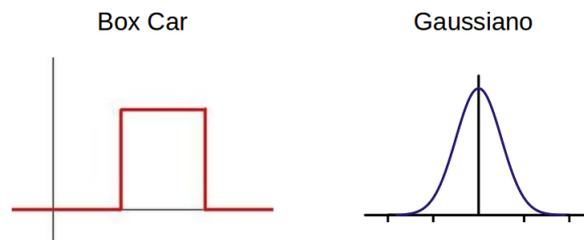
$$f_A(t) = \begin{cases} 4 \operatorname{sen}\left(\frac{2\pi}{3}\left(t + \frac{5}{12}\right)\right) - 3 \operatorname{sen}\left(\frac{2\pi}{6}\left(t - \frac{10}{12}\right)\right) + \epsilon & t = 0, \frac{1}{12}, \frac{2}{12}, \dots, 25 \\ 4 \operatorname{sen}\left(\frac{2\pi}{3}\left(t - \frac{5}{12}\right)\right) - 3 \operatorname{sen}\left(\frac{2\pi}{6}\left(t + \frac{10}{12}\right)\right) + \epsilon & t = 0, \left(25 + \frac{1}{12}\right), \left(25 + \frac{2}{12}\right), \dots, 50 \end{cases} \quad (\text{A.1})$$

$$f_B(t) = \left\{ \operatorname{sen}\left(\frac{2\pi}{3}t\right) + 3 \operatorname{sen}\left(\frac{2\pi}{6}t\right) + \epsilon \quad t = 0, \frac{1}{12}, \frac{2}{12}, \dots, 50 \right. \quad (\text{A.2})$$

Mais detalhes sobre as séries temporais e suavizações estão disponíveis em [Aguiar and Soares \(2014\)](#).

Conforme descrito no Capítulo 3, o processo de suavização pode alterar o resultado das análises da coerência *wavelet* e da fase. Neste estudo, utilizam-se as funções *Box Car* e Gaussiana, apresentadas na Figura A.1, com variações de parâmetros para exemplificar essas diferenças.

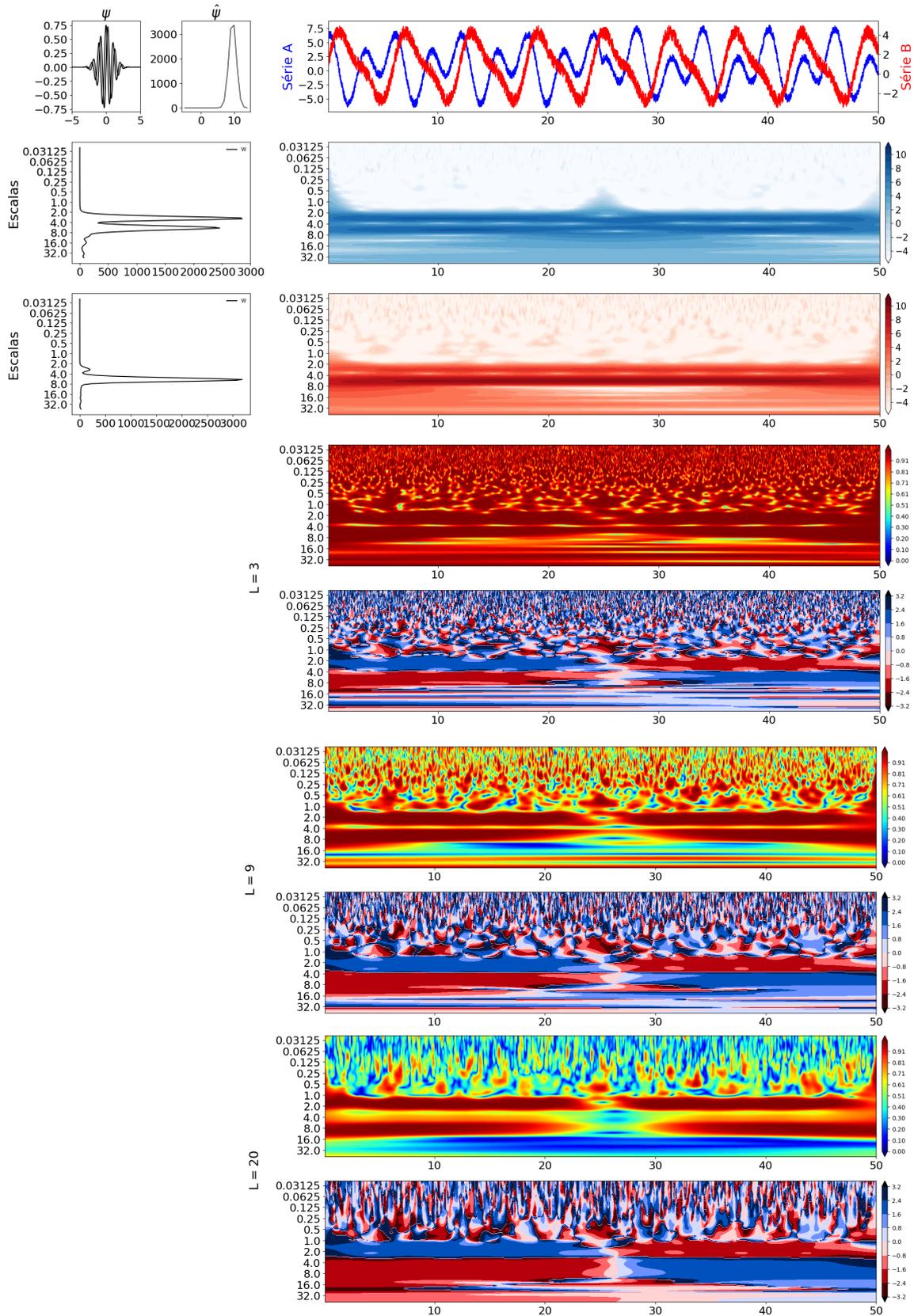
Figura A.1 - Filtros de suavização.



Fonte: Produção do Autor.

No primeiro estudo, utiliza-se a janela *Box Car* com variações de quantidade de valores não nulos (L). A Figura A.2 apresenta a coerência *wavelet* e a fase com variações do parâmetro L.

Figura A.2 - Exemplo de coerência *wavelet* com janela *Box Car*.



Fonte: Produção do Autor.

Primeiramente, nos escalogramas, observa-se que ambos os sinais possuem estruturas entre as escalas que representam 2 anos e 8 anos. Também é possível observar que a variação do parâmetro  $L$  altera intensamente as características dos gráficos de coerência e fase.

Utilizando uma janela de tamanho  $L = 3$ , as características do gráfico de coerência se torna muito ruidosa, dificultando a localização das reais correlações dos sinais. Verifica-se que é possível localizar duas estruturas de alta coerência nas escalas em que os sinais se relacionam; porém, no tempo próximo aos 25 anos, observa-se de forma discreta a quebra da coerência dos sinais. Pelo gráfico da fase é possível observar uma ótima localização em escalas, tornando claro os processos de atrasos e adiantamentos entre os sinais.

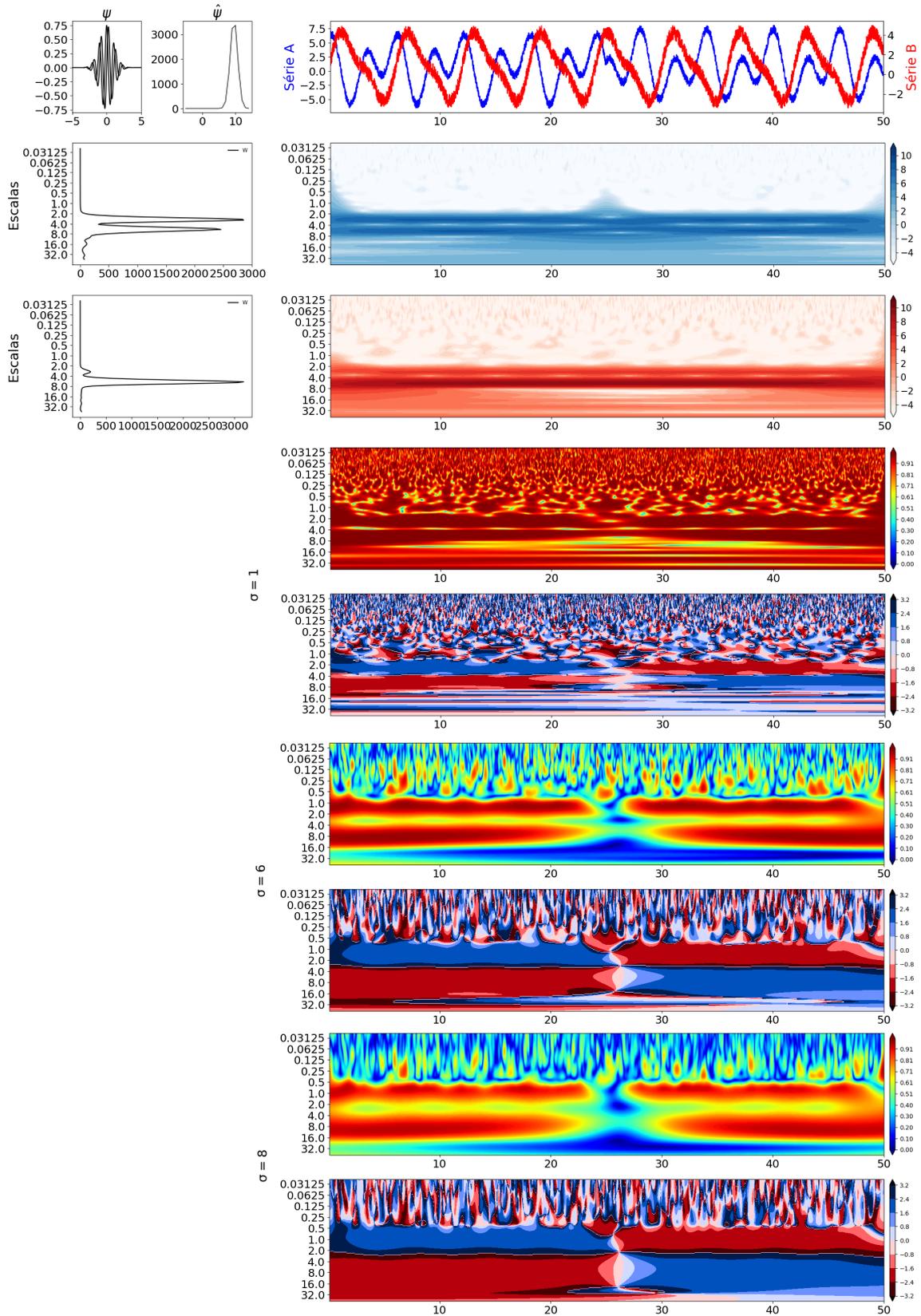
Utilizando o parâmetro  $L = 9$ , observa-se que a estrutura das relações tempo-escala entre os sinais tornam-se mais visíveis. A principal diferença neste caso é um pequeno deslocamento na localização das escalas relacionadas, em que se passa a observar que a localização da quebra de coerência entre os sinais se torna mais clara. Ainda assim observa-se muitas estruturas de relação tempo-escala onde há ruídos. O mesmo fenômeno de deslocamento em escalas observa-se na fase da coerência.

No caso  $L = 20$ , há um maior deslocamento em localização das escalas de relação; porém torna-se muito claro a quebra de coerência entre os sinais no momento 25. As estruturas dos ruídos se tornam menos aparentes.

Em resumo, constata-se que a variação dos parâmetros do filtro de suavização modifica as estruturas de ruído e altera a otimização da localização entre tempo e escalas.

No segundo estudo, utiliza-se a janela Gaussiana com variação de desvio padrão ( $\sigma$ ). A Figura A.3 apresenta a coerência *wavelet* e a fase com variações de  $\sigma$ .

Figura A.3 - Exemplo de coerência *wavelet* com janela *Gaussiana*.



Fonte: Produção do Autor.

De forma geral, observa-se que a alteração de parâmetros do filtro gaussiano destaca as localizações em tempo e escala nos gráficos de coerência e da fase. Devidas à diferença da formulação dos filtros, as alterações acontecem de formas distintas, como esperado.

Como conclusão, entende-se que a escolha de filtros deve ser realizada de acordo com o contexto da análise, por entender seus efeitos nas análises complementares das relações tempo-escala.



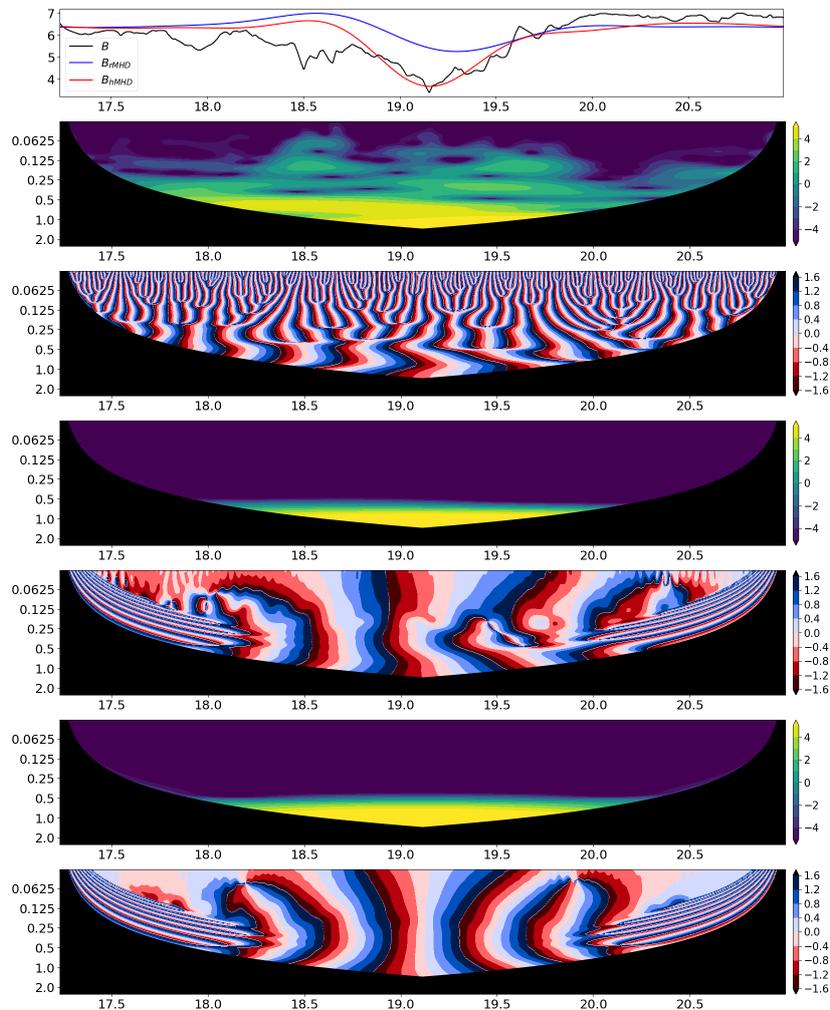
## Apêndice B: Análise complementar dos Dados da Sonda Cassini

Nesta sessão são descritas as análises multiescala e de similaridade do módulo do campo magnético interplanetário e das componentes  $B_x$  e  $B_y$  do estudo comparativo entre dados de campo magnético da Sonda Cassini e provenientes de modelo MHD Resistivo e Hall.

### B.1 Módulo do campo magnético

Apresenta-se na Figura B.1 as características tempo-escala dos dados observacionais e dos modelos rMHD e hMHD para o módulo do campo magnético  $B$ .

Figura B.1 - Dados do módulo do campo magnético  $B$ , escalogramas e fases.



Fonte: Produção do Autor.

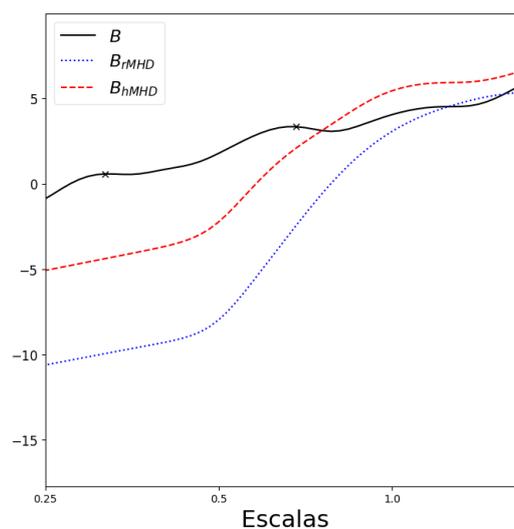
No escalograma do dado observacional é possível verificar estruturas distribuídas em todas as escalas, representando toda a complexidade do sinal. As principais estruturas estão localizadas entre as escalas 0,5 e 2. Observa-se também no gráfico da fase as diversas mudanças de fase ocorrendo nessas escalas.

No escalograma dos dados do modelo rMHD, observam-se estruturas apenas entre as escalas de 0,5 a 2 horas. Percebe-se que o modelo resistivo não possui sensibilidade para reproduzir as escalas menores do que 0,5 horas. As escalas entre 0,5 e 1 não foram reproduzidas com precisão, apresentando uma forte componente de variação nessas escalas. No gráfico da fase localiza-se claramente as mudanças de fase entre os momentos 18,5 e 19,5.

Analisando o escalograma dos dados do modelo hMHD, percebe-se também que não foram apresentadas as componentes de variações rápidas. Entre as escalas 0,5 e 1 hora, verifica-se que o modelo de Hall representou de forma muito mais precisa as energias dessas estruturas. No gráfico da fase, localiza-se a variação de fase em momento mais preciso nas escalas de maiores energias em relação ao modelo resistivo.

Para analisar globalmente a distribuição de energia dos sinais, analisa-se o espectro global. A Figura B.2 apresenta a distribuição de energias globais em cada escala.

Figura B.2 - Espectro wavelet global para o módulo do campo magnético  $B$  referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado).

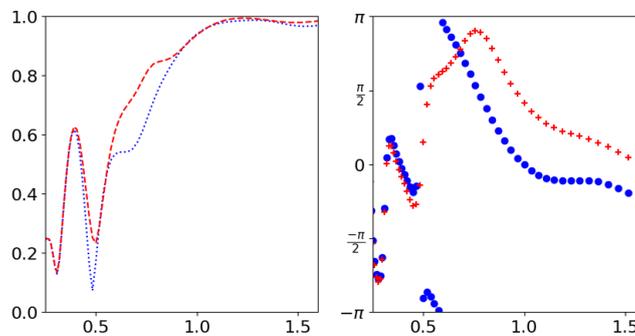


Fonte: Produção do Autor.

No espectro global, observa-se que a curva que representa o dado observacional possui maior quantidade de marcações. Observa-se que, em ambos os modelos, obtiveram-se marcações de máximos próximas às escalas 1,5 e 2,2, onde há correspondência com o dado observacional.

Na Figura B.3 estão apresentados as correlações por escala (módulo à esquerda e fase à direita) entre os dados observacionais de  $B$  e os obtidos pelos modelos rMHD (curva em azul) e hMHD (curva em vermelho).

Figura B.3 - Módulo e fase da correlação por escala de  $B$  em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul, pontilhado) e hMHD (vermelho, tracejado).

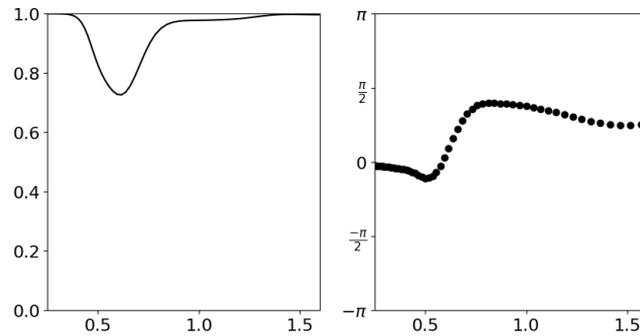


Fonte: Produção do Autor.

Observa-se excelente correlação de ambos os modelos a partir da escala de 1 hora, em que a curva do modelo hMHD apresenta valores ligeiramente maiores. Observa-se também que ambas as curvas são muito parecidas, destaca-se um leve deslocamento de amplitude de correlação nas escalas próximas a 0,5 horas. A curva do modelo hMHD apresenta maior em fase nessas escalas, com tendência de quadratura; já o modelo rMHD apresenta menos escalas em fase e uma maior tendência de quadratura.

Na Figura B.4 estão apresentadas as correlações por escala (módulo à esquerda e fase à direita) entre os dados simulados de  $B$  pelos modelos rMHD e hMHD.

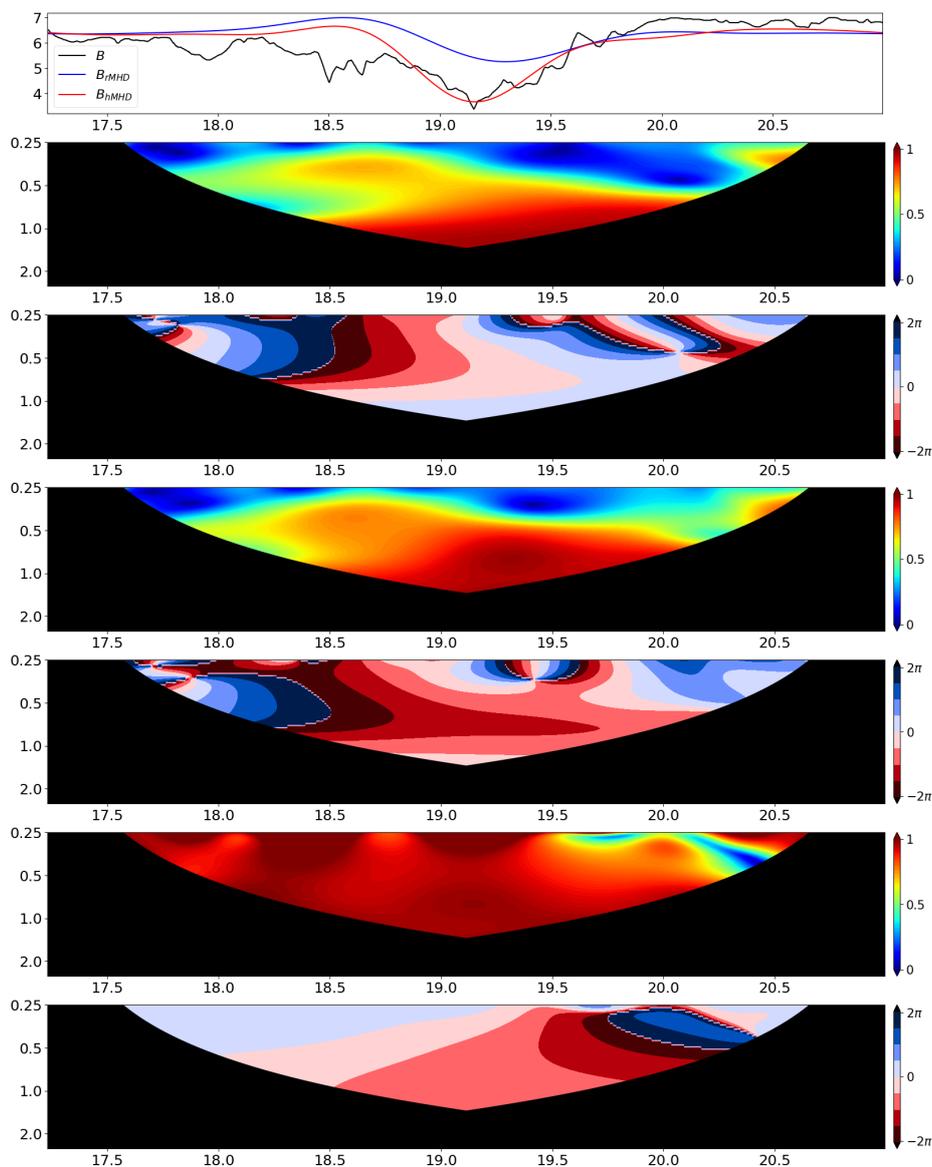
Figura B.4 - Módulo e fase da correlação por escala dos dados modelados pelo rMHD e hMHD.



Fonte: Produção do Autor.

Nas escalas menores que 0,25, a alta correlação próxima a 1 se dá pela incapacidade de ambos os modelos de reproduzir os dados observacionais, e obtendo energias praticamente constantes, como pode ser observado nos escalogramas (Figura B.1). Entre as escalas 1 – 1,5, a correlação por escalas de ambos é verdadeiramente a máxima próxima a 1. Nesse caso também é apresentada uma quadratura na fase dessas escalas.

Figura B.5 - Dados de  $B$ , coerência e diferença de fase.



Fonte: Produção do Autor.

Nos gráficos de coerência *wavelet*, confirmam-se as relações observadas na correlação por escalas. Comparando os dois primeiros gráficos de coerência, observa-se que a relação dos dados observacionais e do modelo hMHD possuem maior coerência nas escalas onde há correlações.

Na diferença de fase, observa-se com precisão a relação de quadratura entre os sinais. Observa-se na relação com o modelo rMHD que há um leve adiantamento nas escalas próximas a 1 hora. No modelo hMHD, esse adiantamento é mais intenso, e

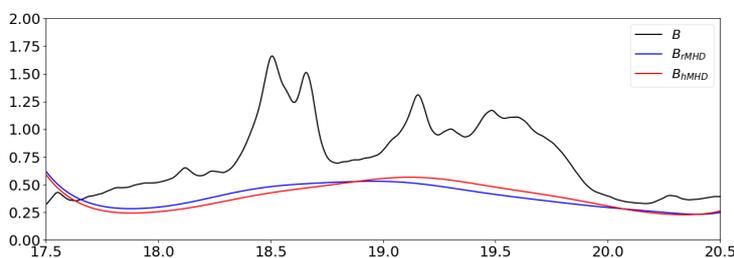
persiste em quase todas as escalas.

A coerência entre os dados de ambos os modelos apresenta forte coerência nas escalas de variações mais rápidas, devido ao fato já apresentado anteriormente nas correlações por escala; já nas escalas onde de fato há informações, observa-se que há uma menor coerência entre os modelos, visto que o modelo resistivo não representou com precisão as escalas entre 0,5 e 1 hora.

Observa-se na diferença de fase entre os dados modelados que há forte adiantamento do modelo de Hall, a relação de quadratura é apresentada com mais intensidade devido aos atrasos já mencionados do modelo resistivo.

Analisa-se a entropia dos sinais na Figura B.6. Este gráfico da entropia confirma que os dados simulados não conseguem acompanhar todas a complexidade dos dados observados; por outro lado, ele consegue dar indicativos de onde se encontram as maiores complexidades, neste caso entre 18 – 19,5 horas.

Figura B.6 - Dados de  $B$  - Entropia Wavelet.



Fonte: Produção do Autor.

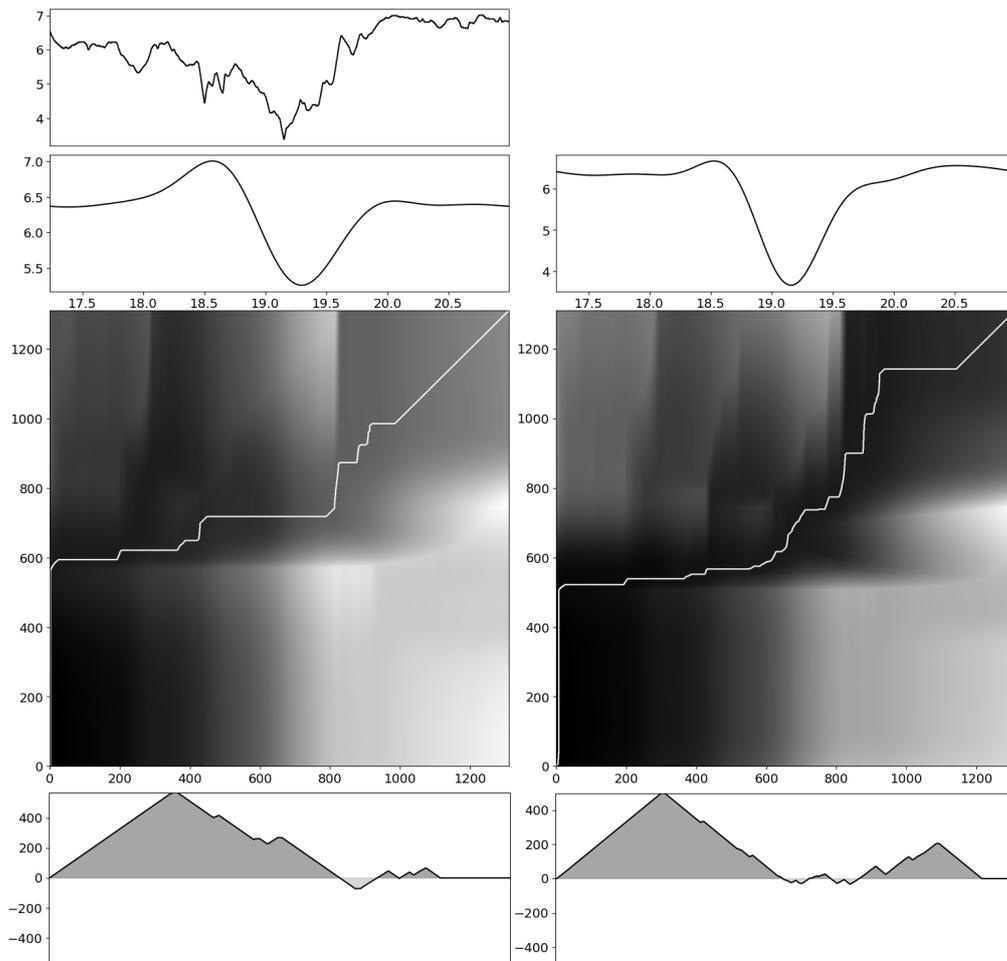
### B.1.1 Análise das métricas de similaridade entre dados observacionais e dados simulados

No cenário  $B \times B_{rMHD}$ , obtiveram-se distância Euclidiana 32,85, DTW 0,22 e RMSE 0,91. Já no cenário  $B \times B_{hMHD}$ , obtiveram-se distância Euclidiana 21,3, DTW 0,1 e RMSE 0,59. Observa-se que os valores para as métricas obtidos pelo modelo hMHD foram melhores.

Os melhores valores para as métricas no cenário  $B \times B_{hMHD}$  são motivados pela melhor reprodução do modelo de Hall em mais escalas em comparação ao modelo resistivo, além de possuir amplitude mais próxima ao dado observado.

Para avaliar os deslocamentos realizados na DTW, analisam-se os deslocamentos na Figura B.7.

Figura B.7 - Comparativo de matrizes de distância e caminhos  $\mathfrak{W}$  entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita).



Fonte: Produção do Autor.

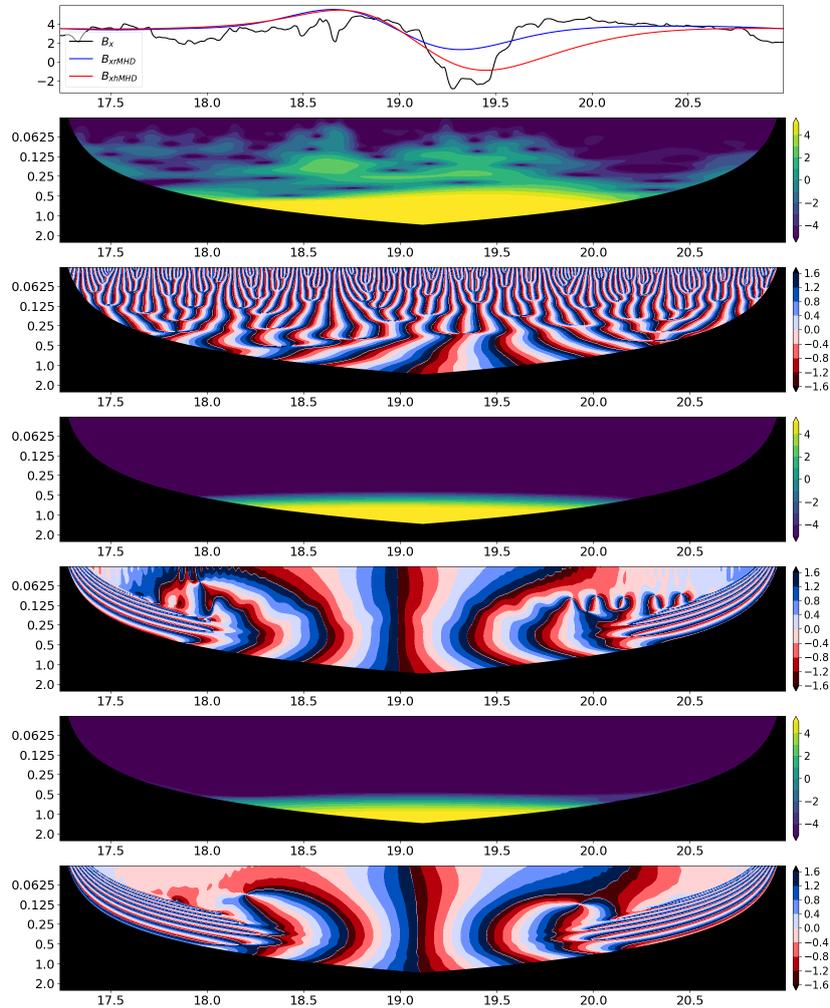
No primeiro cenário, obtiveram-se as áreas de 6923 pontos de atrasos e 399212 pontos de adiantamentos para o caminho  $\mathfrak{W}$ . No segundo cenário obteve-se área de 4269 pontos de atrasos e 328858 pontos de adiantamentos.

## B.2 Componente magnética na direção x

Nesta subseção, analisam-se os dados da componente  $B_x$  do campo magnético. Analisam-se as características tempo-escala dos dados observacionais e dos modelos

rMHD e hMHD na Figura B.8.

Figura B.8 - Dados da componente  $B_x$ , escalogramas e fases.



Fonte: Produção do Autor.

No escalograma do dado observacional é possível verificar estruturas distribuídas em todas as escalas, representando toda a complexidade do sinal. As principais estruturas estão localizadas entre as escalas 0,25 e 1,5. Observam-se também no gráfico da fase as diversas mudanças de fase ocorrendo nessas escalas.

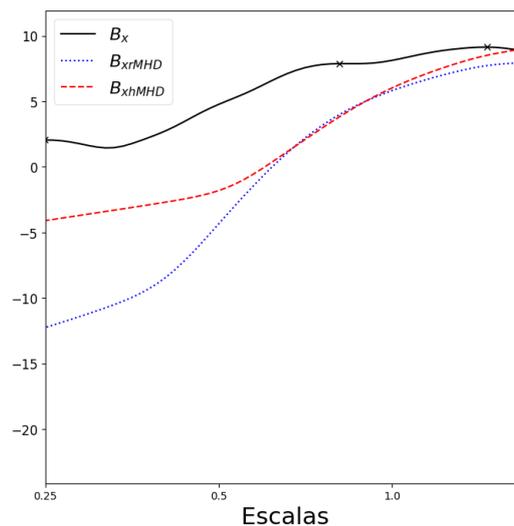
No escalograma dos dados do modelo rMHD, observam-se estruturas apenas entre as escalas de 0,5 a 1,5 horas. Percebe-se que o modelo resistivo não possui sensibilidade para reproduzir as escalas menores do que 0,5 horas. As escalas entre 0,5 e 1 não foram reproduzidas com precisão, apresentando uma forte componente de variação

nessas escalas. No gráfico da fase, localizam-se claramente as mudanças de fase entre os momentos 18,5 e 19,5.

Analisando o escalograma dos dados do modelo hMHD, percebe-se que também não foram apresentadas as componentes de variações rápidas. Entre as escalas 0,5 e 1 hora, verifica-se que o modelo de Hall representou de forma muito mais precisa as energias dessas estruturas. No gráfico da fase localiza-se a variação de fase em momento mais preciso nas escalas de maiores energias em relação ao modelo resistivo.

Para analisar globalmente a distribuição de energia dos sinais, analisa-se o espectro global. A Figura B.9 apresenta a distribuição de energias globais em cada escala.

Figura B.9 - Espectro wavelet global para componente  $B_x$  referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado).



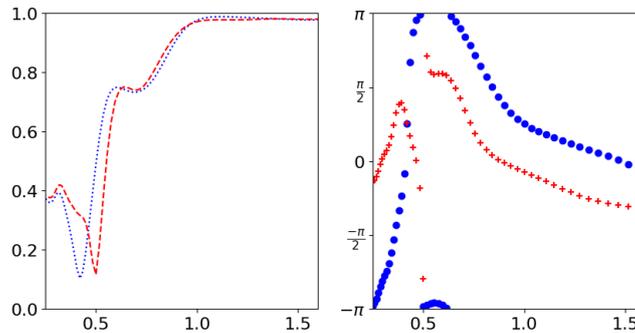
Fonte: Produção do Autor.

No espectro global, observa-se que a curva que representa o dado observacional possui maior quantidade de marcações. Na curva dos dados do modelo rMHD, observa-se que o pico na escala de 1,5 horas não possui correspondência na curva dos dados observacionais. A curva dos dados do modelo hMHD apresenta maior energia no pico da escala de 1,5 horas..

Na Figura B.10 estão apresentados as correlações por escala (módulo à esquerda e fase à direita) entre os dados observacionais de  $B_x$  e os obtidos pelos modelos rMHD

(curva em azul) e hMHD (curva em vermelho).

Figura B.10 - Módulo e fase da correlação por escala da componente  $B_x$  em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul, pontilhado) e hMHD (vermelho, tracejado).

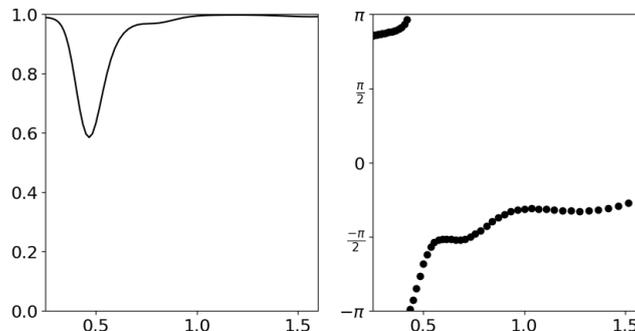


Fonte: Produção do Autor.

Observa-se excelente correlação de ambos os modelos a partir da escala de 1 hora. Observa-se também que ambas as curvas são muito parecidas, destaca-se um leve deslocamento de amplitude de correlação nas escalas próximas a 0,5 horas. Em ambas as curvas, nas maiores escalas, há a tendência de comportamento em fase. Enquanto, nas escalas menores, o modelo hMHD apresenta maiores escalas com relações mais uniformes; no modelo rMHD, há mais escalas em quadratura.

Na Figura B.11 está apresentada a correlações por escala (módulo à esquerda e fase à direita) entre os dados simulados de  $B_x$  pelos modelos rMHD e hMHD.

Figura B.11 - Módulo e fase da correlação por escala dos dados modelados pelo rMHD e hMHD.

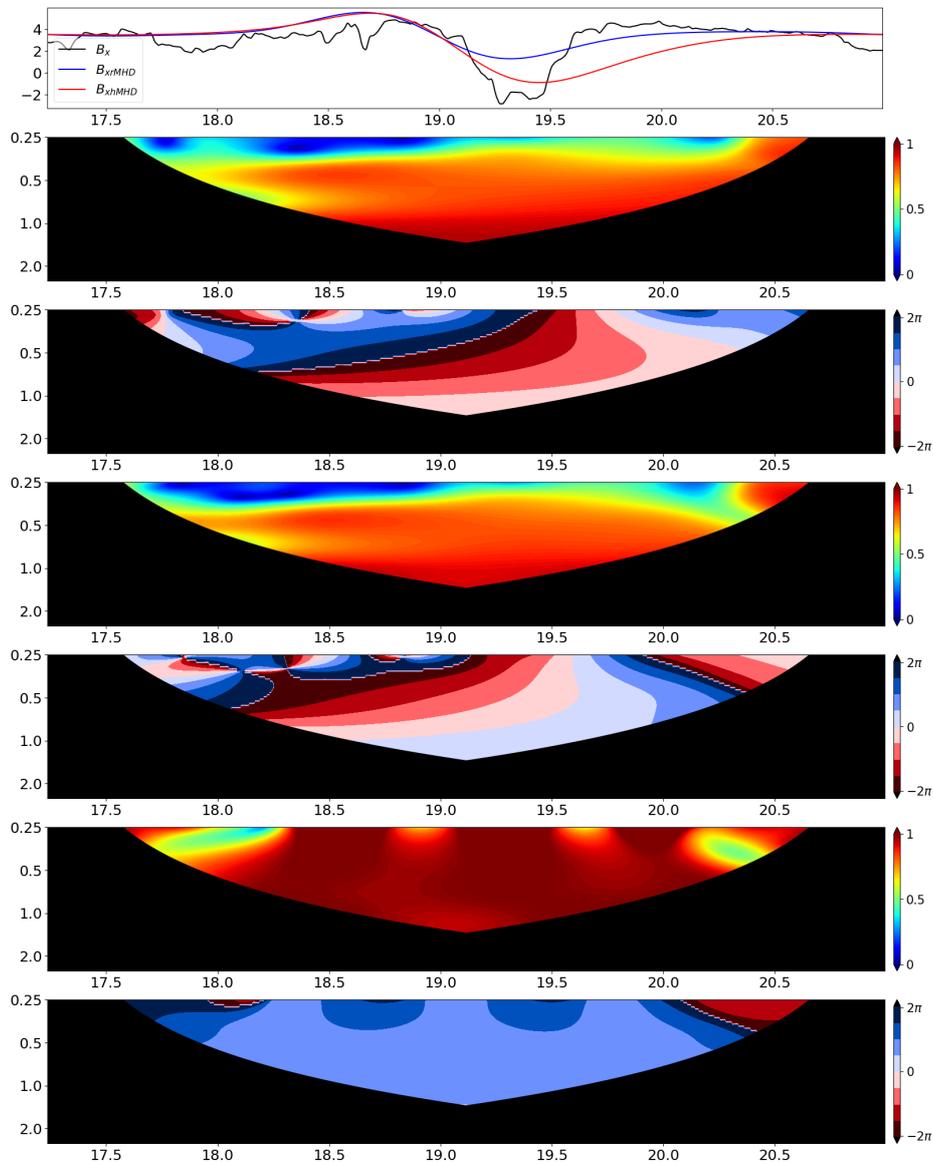


Fonte: Produção do Autor.

Nas escalas menores que 0,25, a alta correlação próxima a 1 se dá pela incapacidade de ambos os modelos de reproduzir os dados observacionais, e obtendo energias praticamente constantes, como pode ser observado nos escalogramas (Figura B.8). Entre as escalas 1–2, a correlação por escalas de ambos é verdadeiramente a máxima, próxima a 1. Nesse caso é apresentado uma tendência de comportamento em fase entre os sinais modelados.

Apresenta-se os gráficos de coerência e diferença de fase na Figura B.12.

Figura B.12 - Dados da componente  $B_x$ , coerência e diferença de fase.



Fonte: Produção do Autor.

Nos gráficos de coerência *wavelet*, confirmam-se as relações observadas na correlação por escalas. Comparando os dois primeiros gráficos de coerência, observa-se que a relação dos dados observacionais e do modelo hMHD possuem maior coerência nas escalas onde há correlações.

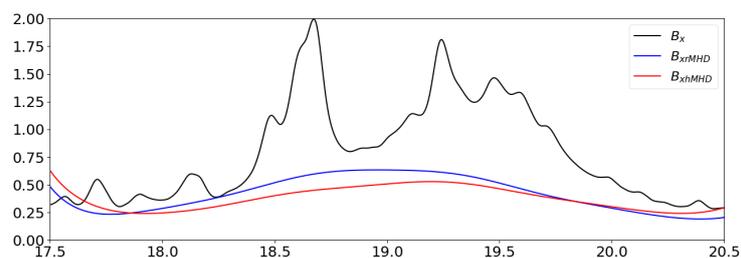
Na diferença de fase observa-se com precisão a relação de quadratura entre os sinais. A relação entre os dados observados e os do modelo resistivo aponta que o dado simulado está atrasado entre as escalas de 1 a 1,5 horas. Essa relação é inversa nos dados simulados do modelo de Hall, apontando adiantamento nessas mesmas escalas. Porém essas relações de atraso e adiantamento possuem baixo valor, comprovando a tendência de comportamento em fase.

A coerência entre os dados de ambos os modelos apresenta forte coerência nas escalas de variações mais rápidas, devido ao fato já apresentado anteriormente nas correlações por escala, já, nas escalas onde de fato há informações, observa-se que há uma menor coerência entre os modelos, visto que o modelo resistivo não representou com precisão as escalas entre 0,5 e 1 hora.

Observa-se na diferença de fase entre os dados modelados que há forte adiantamento do modelo de Hall nas menores escalas. Já nas maiores escalas a relação se inverte, havendo adiantamento no sinal do modelo resistivo.

Analisa-se a entropia dos sinais na Figura B.13.

Figura B.13 - Componente  $B_x$  - Entropia Wavelet.



Fonte: Produção do Autor.

O gráfico da entropia (Figura B.13) confirma que os dados simulados não conseguem acompanhar todas as complexidades dos dados observados. Por outro lado, ele consegue dar indicativos de onde se encontram as maiores complexidades, neste caso entre 18 – 19,5 horas.

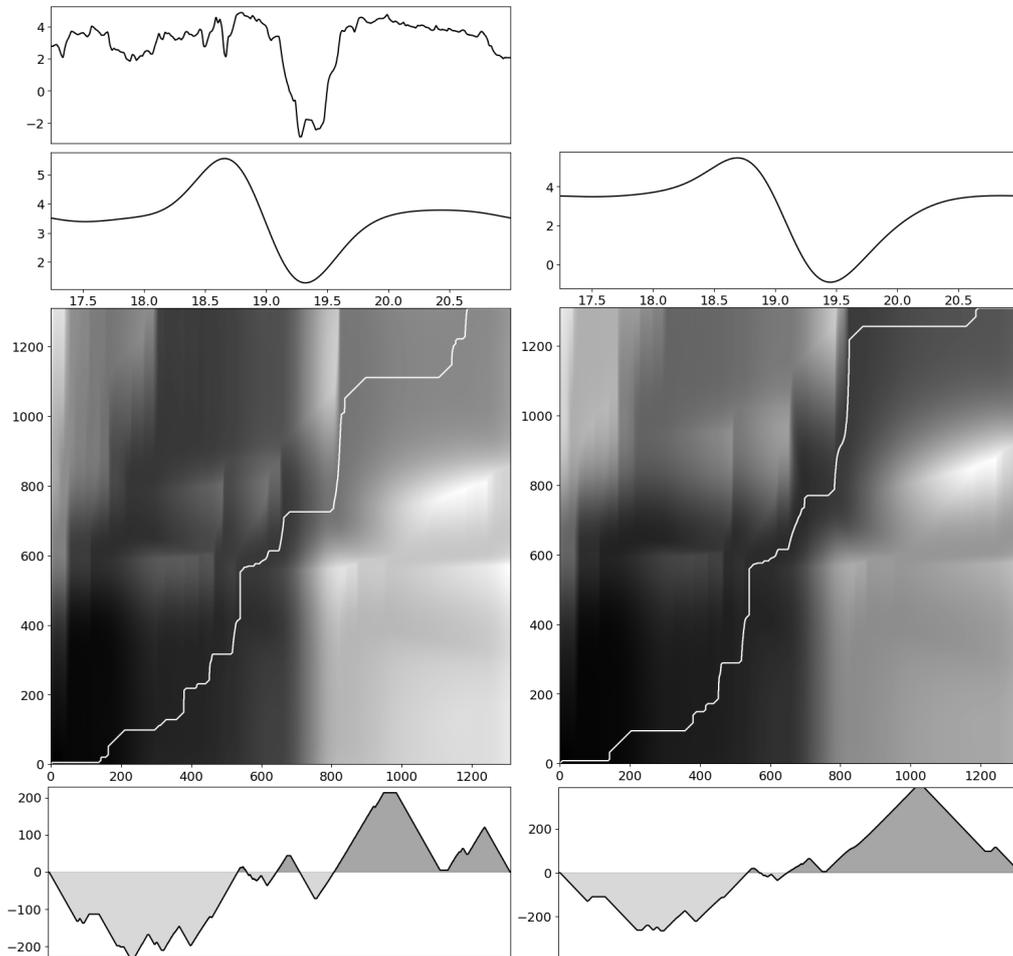
### B.2.1 Análise das métricas de similaridade entre dados observacionais e dados simulados

No cenário  $B_x \times B_{x,r\text{MHD}}$ , obtiveram-se distância Euclidiana 42,77, DTW 0,34 e RMSE 1,32. Já no cenário  $B_x \times B_{x,h\text{MHD}}$ , obtiveram-se distância Euclidiana 57,29, DTW 0,28 e RMSE 1,58. Observa-se que os valores para as métricas obtidos pelo modelo hMHD foram melhores.

Os melhores valores para as métricas no cenário  $B_x \times B_{x,h\text{MHD}}$  são motivados pela melhor reprodução do modelo de Hall em mais escalas em comparação ao modelo resistivo, além de possuir amplitude mais próxima ao dado observado.

Para avaliar os deslocamentos realizados na DTW analisam-se os deslocamentos na Figura B.14.

Figura B.14 - Comparativo de matrizes de distância e caminhos  $\mathfrak{W}$  entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita).



Fonte: Produção do Autor.

No primeiro cenário, obtiveram-se áreas de 133611 pontos de atrasos e 82482 pontos de adiantamentos para o caminho  $\mathfrak{W}$ . No segundo cenário, obtiveram-se área de 142437 pontos de atrasos e 188072 pontos de adiantamentos.

Analisando as métricas de similaridade, observa-se que, ponto a ponto, o modelo resistivo reproduziu uma série mais similar ao dado observado; porém ao se aplicar a DTW, que lida com a ocorrência da distorção temporal, obteve-se um valor mais baixo.

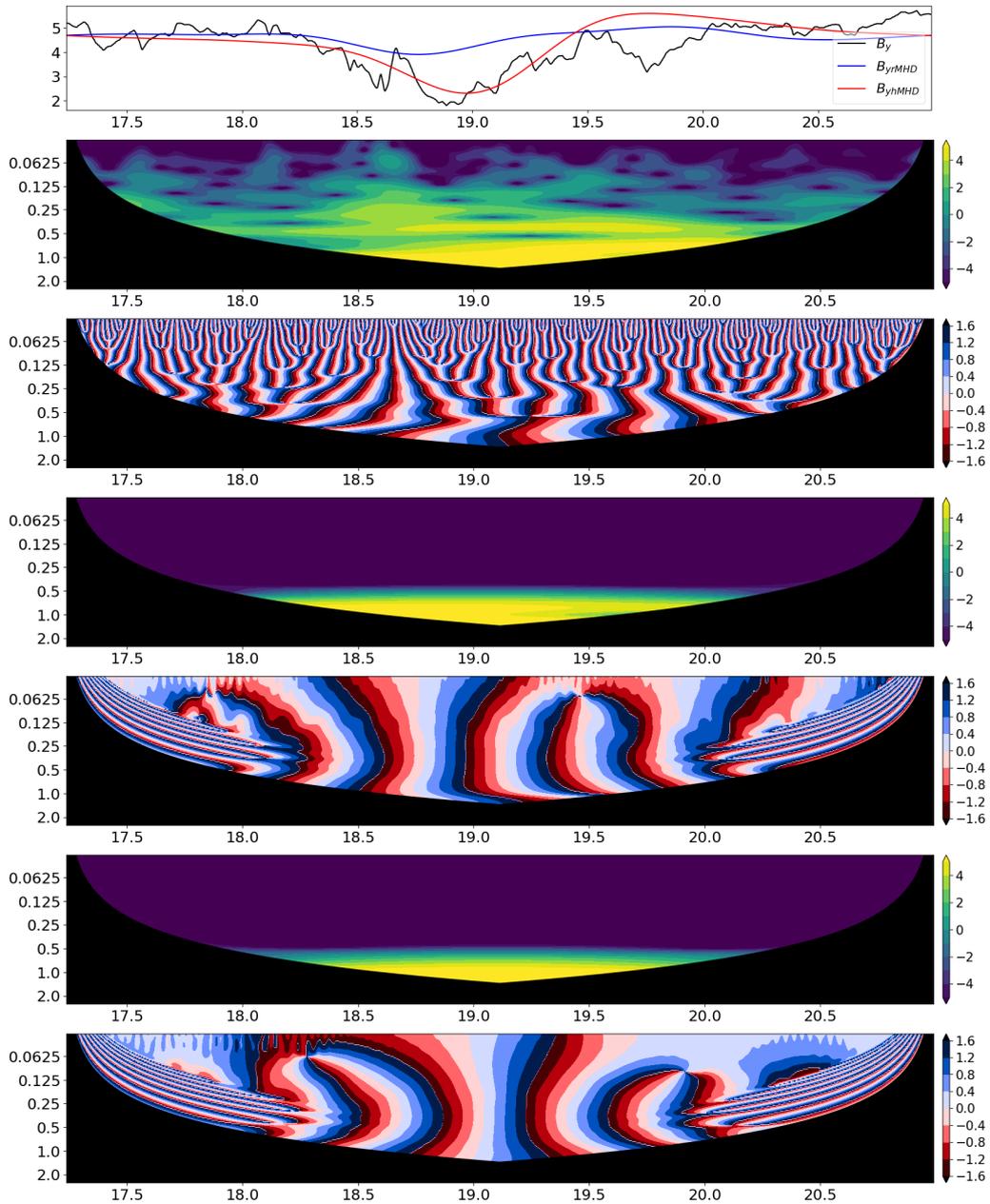
### B.3 Componente magnética na direção $y$

Nesta subseção, analisam-se os dados da componente  $B_y$  do campo magnético. O comportamento tempo-escala dos dados observacionais e dos modelos rMHD e hMHD é apresentado na Figura B.15. No escalograma do dado observacional é possível verificar estruturas distribuídas em todas as escalas, representando toda a complexidade do sinal. As principais estruturas estão localizadas entre as escalas 0,25 e 1,5. Observa-se também no gráfico da fase as diversas mudanças de fase ocorrendo nessas escalas.

No escalograma dos dados do modelo rMHD observa-se estruturas apenas entre as escalas de 0,5 a 1,5 horas, percebe-se que o modelo resistivo não possui sensibilidade para reproduzir as escalas menores do que 0,5 horas, as escalas entre 0,5 e 1 não foram reproduzidas com precisão, apresentando uma forte componente de variação nessas escalas. No gráfico da fase, localiza-se claramente as mudanças de fase entre os momentos 18,5 e 19,5.

Analisando o escalograma dos dados do modelo hMHD, percebe-se que também não foram apresentadas as componentes de variações rápidas. Entre as escalas 0,5 e 1 hora verifica-se que o modelo de Hall representou de forma muito mais precisa as energias dessas estruturas. No gráfico da fase, localiza-se a variação de fase em momento mais preciso nas escalas de maiores energias em relação ao modelo resistivo.

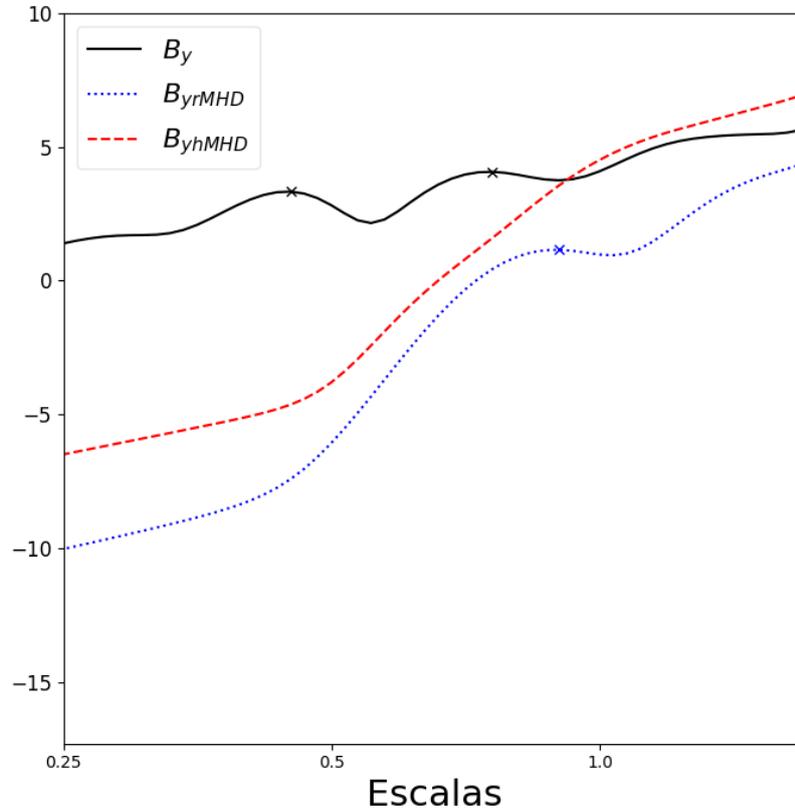
Figura B.15 - Dados da componente  $B_y$ , escalogramas e fases.



Fonte: Produção do Autor.

Para analisar globalmente a distribuição de energia dos sinais, analisa-se o espectro global. A Figura B.16 apresenta a distribuição de energias globais em cada escala.

Figura B.16 - Espectro wavelet global para componente  $B_y$  referente aos dados observacionais (preto) e seus respectivos dados simulados pelos modelos rMHD (azul pontilhado) e hMHD (vermelho tracejado).

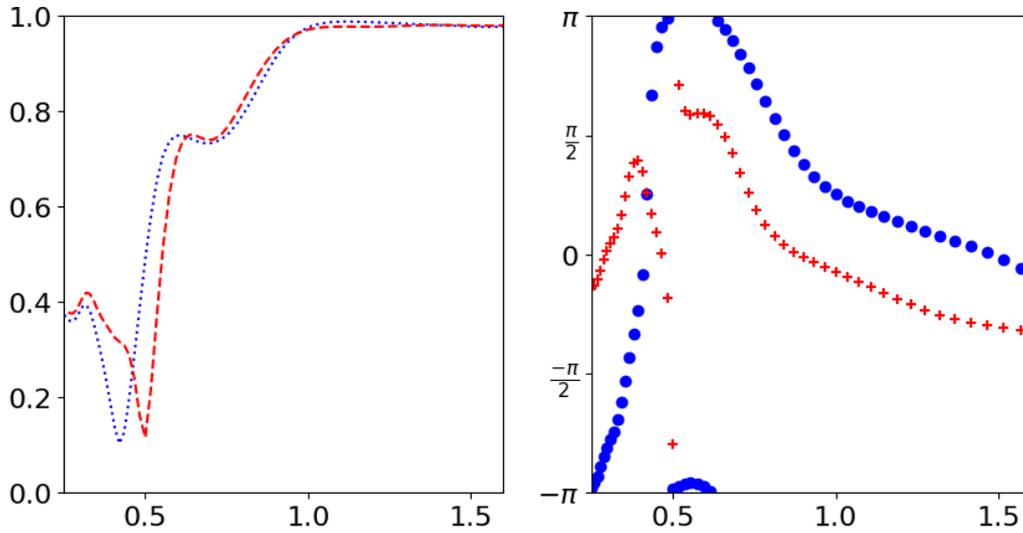


Fonte: Produção do Autor.

No espectro global estão marcados os principais picos de energia nas escalas. Observa-se que a curva que representa os dados observacionais possui maior quantidade de marcações. Na curva dos dados do modelo rMHD, observa-se que o pico na escala de 1 hora não possui correspondência na curva dos dados observacionais. A curva dos dados do modelo hMHD apresentam as mesmas marcações de máximos, com diferença na energia que é superior aos dados observados.

Na Figura B.17 estão apresentados as correlações por escala (módulo à esquerda e fase à direita) entre os dados observacionais de  $B_z$  e os obtidos pelos modelos rMHD (curva em azul) e hMHD (curva em vermelho).

Figura B.17 - Módulo e fase da correlação por escala da componente  $B_y$  em relação ao respectivo dados simulados dessa componente nos modelos rMHD (azul, pontilhado) e hMHD (vermelho, tracejado).

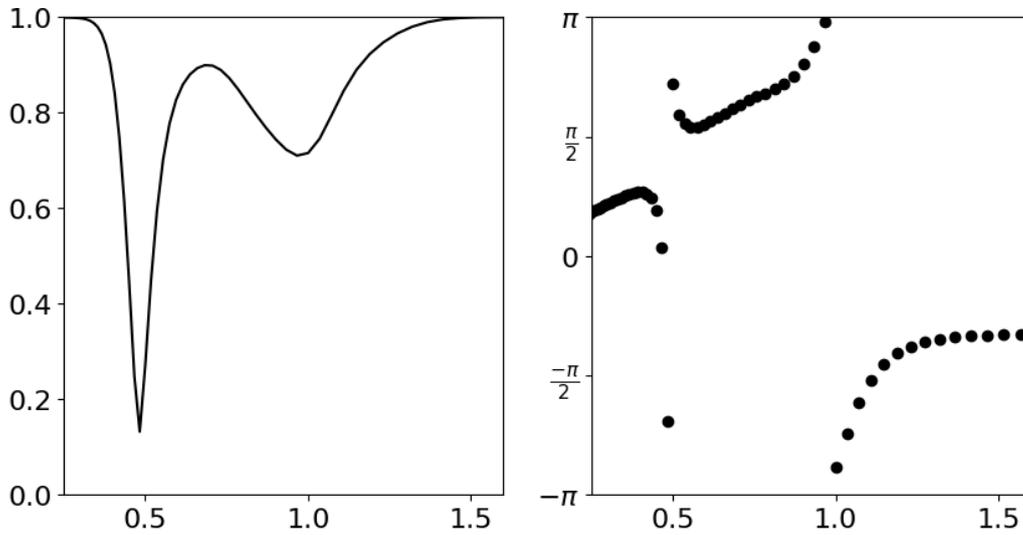


Fonte: Produção do Autor.

Observa-se excelente correlação de ambos os modelos a partir da escala de 0,8 horas. Observa-se também que ambas as curvas são muito parecidas, destaca-se um leve deslocamento de correlação nas escalas próximas a 0,5 horas. Em ambas as curvas, nas maiores escalas há a tendência de comportamento de leve quadratura, já nas escalas menores o modelo hMHD apresenta maiores escalas com relações mais uniformes, já o modelo rMHD há mais escalas em quadratura.

Na Figura B.18 está apresentada a correlações por escala (módulo à esquerda e fase à direita) entre os dados simulados de  $B_y$  pelos modelos rMHD e hMHD.

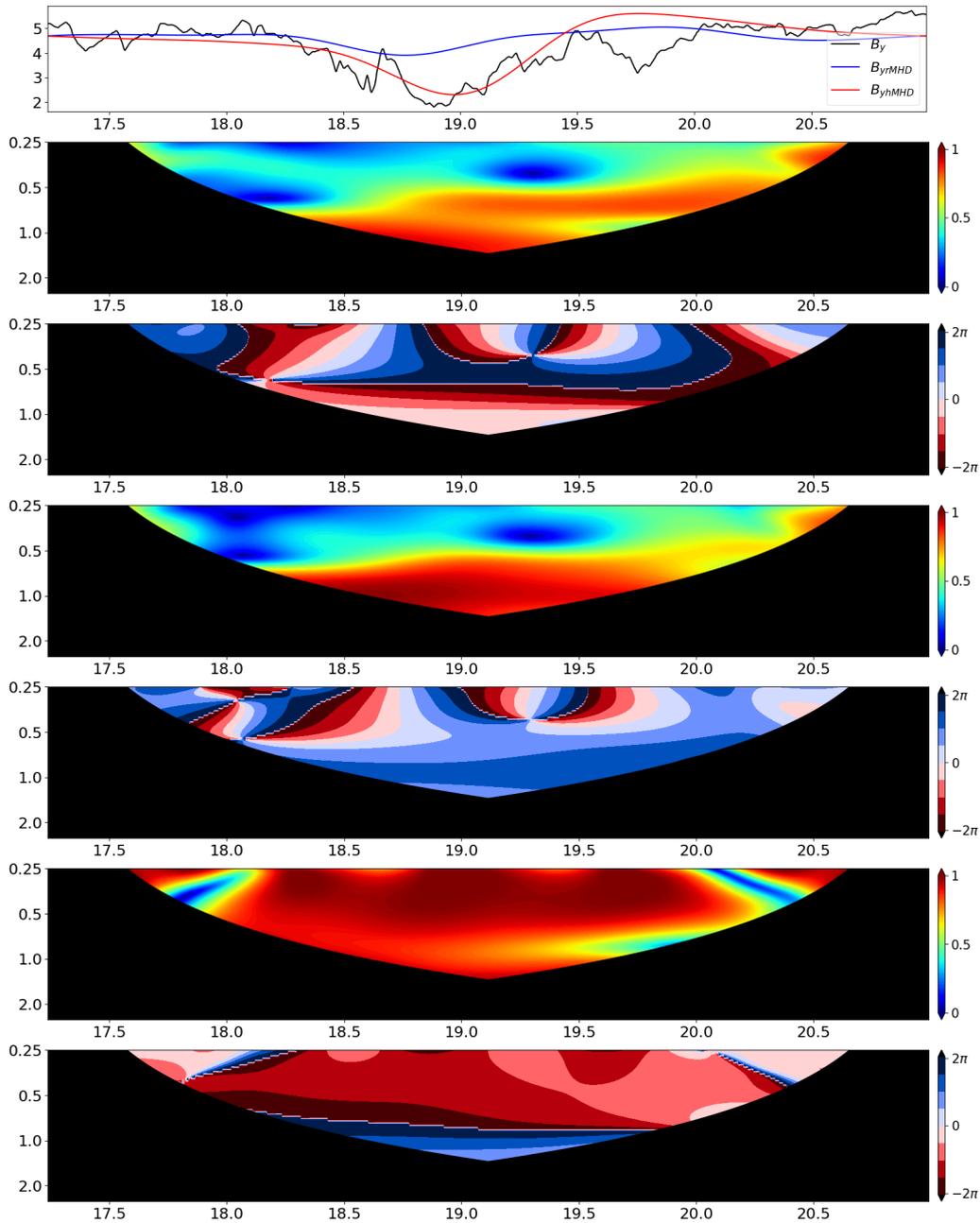
Figura B.18 - Módulo e fase da correlação por escala dos dados modelados pelo rMHD e hMHD.



Fonte: Produção do Autor.

Nas escalas menores que 0,25, a alta correlação próxima a 1 se dá pela incapacidade de ambos os modelos de reproduzir os dados observacionais, e obtendo energias praticamente constantes, como pode ser observado nos escalogramas (Figura B.15). Entre as escalas 1 – 1,5 a correlação por escalas de ambos é verdadeiramente a máxima próxima a 1. Nesse caso também é apresentado uma quadratura com tendência a comportamento em fase nessas escala.

Figura B.19 - Dados da componente  $B_y$ , coerência e diferença de fase.



Fonte: Produção do Autor.

Nos gráficos de coerência *wavelet* confirmam-se as relações observadas na correlação por escalas. Comparando os dois primeiros gráficos de coerência, observa-se que a relação dos dados observacionais e do modelo hMHD possuem maior coerência nas escalas onde há correlações.

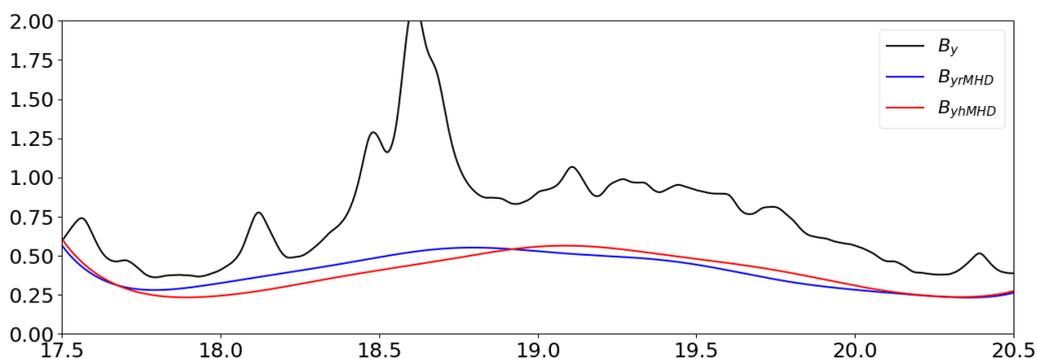
Na diferença de fase, observa-se com precisão a relação de quadratura entre os sinais. A relação entre os dados observados e os do modelo resistivo aponta que o dado simulado está atrasado entre as escalas de 1 a 1,5 horas. Essa relação é inversa nos dados simulados do modelo de Hall, apontando adiantamento nessas mesmas escalas. Nas escalas menores é possível observar que ambos os modelos estão adiantados em relação ao dado observado na maior parte do tempo.

A coerência entre os dados de ambos os modelos apresenta forte coerência nas escalas de variações mais rápidas, devido ao fato já apresentado anteriormente nas correlações por escala, já nas escalas onde de fato há informações, observa-se que há uma menor coerência entre os modelos, visto que o modelo resistivo não representou com precisão as escalas entre 0,5 e 1 hora.

Observa-se na diferença de fase entre os dados modelados que há forte adiantamento do modelo de Hall, a relação de quadratura é apresentada com mais intensidade devido aos atrasos já mencionados do modelo resistivo.

Apresenta-se a entropia *wavelet* da componente  $B_y$  na Figura B.20. O gráfico da entropia (Figura B.20) confirma que os dados simulados não conseguem acompanhar todas a complexidade dos dados observados, por outro lado, ele consegue dar indicativos de onde se encontram as maiores complexidades, neste caso entre 18 – 19,5 horas.

Figura B.20 - Componente  $B_y$  - Entropia *wavelet*.



Fonte: Produção do Autor.

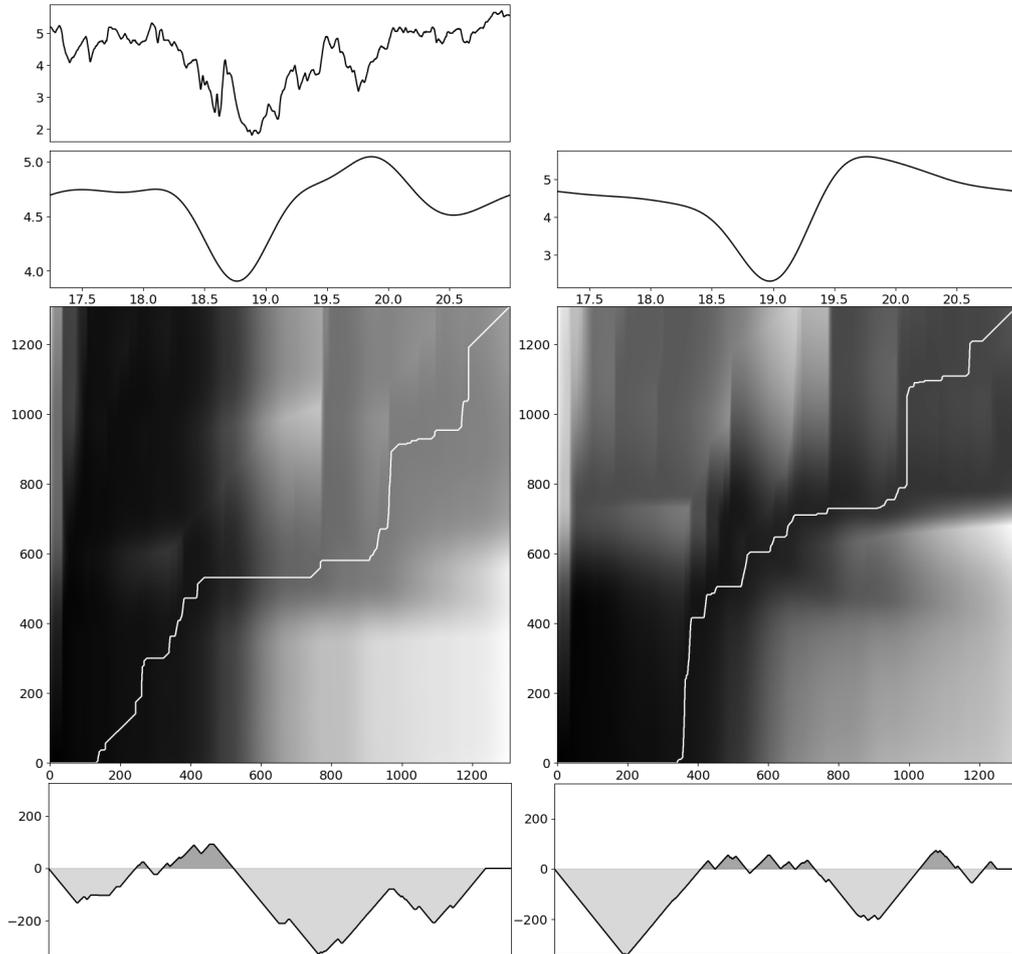
### B.3.1 Análise das métricas de similaridade entre dados observacionais e dados simulados

No cenário  $B_y \times B_{y,rMHD}$ , obtiveram-se distância Euclidiana 31,32, DTW 0,22 e RMSE 0,91. Já no cenário  $B_y \times B_{y,hMHD}$  obtiveram-se distância Euclidiana 24,4, DTW 0,13 e RMSE 0,67. Observa-se que os valores para as métricas obtidos pelo modelo hMHD foram melhores.

Os melhores valores para as métricas no cenário  $B_y \times B_{y,hMHD}$  são motivados pela melhor reprodução do modelo de Hall em mais escalas em comparação ao modelo resistivo, além de possuir amplitude mais próxima ao dado observado.

Para avaliar os deslocamentos realizados na DTW analisam-se os deslocamentos na Figura B.21.

Figura B.21 - Comparativo de matrizes de distância e caminhos  $\mathfrak{W}$  entre dados observacionais e dos modelos rMHD (a esquerda) e hMHD (a direita).



Fonte: Produção do Autor.

No primeiro cenário, obtiveram-se áreas de 226927 pontos de atrasos e 17647 pontos de adiantamentos para o caminho  $\mathfrak{W}$ . No segundo cenário obtiveram-se áreas de 188402 pontos de atrasos e 21457 pontos de adiantamentos.

#### B.4 Resumo dos valores das métricas de similaridade

A Tabela B.1 apresenta uma compilação das métricas de similaridade.

Tabela B.1 - Similaridades entre os dados observacionais e dos modelos rMHD e hMHD referentes ao campo magnético.

Séries	Caminho W (área)		Distância		RMSE
	Atrasos	Adiantamentos	Euclidiana	DTW	
$[ \mathbf{B} ,  \mathbf{B} _{r\text{MHD}}]$	6.923	399.212	32,85	0,22	0,91
$[ \mathbf{B} ,  \mathbf{B} _{h\text{MHD}}]$	4.269	328.858	21.30	0,1	0,59
$[B_x, B_{x,r\text{MHD}}]$	133.611	82.482	47.77	0,34	1,32
$[B_x, B_{x,h\text{MHD}}]$	142.437	188.072	57,29	0,28	1,58
$[B_y, B_{y,r\text{MHD}}]$	226.927	17.647	31,32	0,22	0,87
$[B_y, B_{y,h\text{MHD}}]$	188.402	21.457	24,4	0,15	0,67
$[B_z, B_{z,r\text{MHD}}]$	122.445	223.494	35,6	0,20	0,98
$[B_z, B_{z,h\text{MHD}}]$	44.202	190.446	24,7	0,13	0,68

Fonte: Produção do Autor.

Na análise de similaridade das outras componentes e do modulo do campo magnético obteve-se resultados similares aos analisados na componente  $B_z$ , com exceção a componente  $B_x$ , onde foram obtidos valores maiores nas métricas de similaridade para o modelo hMHD.

## PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

### **Teses e Dissertações (TDI)**

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

### **Manuais Técnicos (MAN)**

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

### **Notas Técnico-Científicas (NTC)**

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

### **Relatórios de Pesquisa (RPQ)**

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

### **Propostas e Relatórios de Projetos (PRP)**

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

### **Publicações Didáticas (PUD)**

Incluem apostilas, notas de aula e manuais didáticos.

### **Publicações Seriadas**

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

### **Programas de Computador (PDC)**

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

### **Pré-publicações (PRE)**

Todos os artigos publicados em periódicos, anais e como capítulos de livros.