



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-14670-TDI/1223

**ESTUDO DA VARIABILIDADE DE SÉRIES TEMPORAIS
AMBIENTAIS ATRAVÉS DE ESPECTROS DE ASSIMETRIA
IMPLEMENTADOS EM SOFTWARE CIENTÍFICO LIVRE**

Eduardo Batista de Moraes Barbosa

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelos Drs. Reinaldo Roberto Rosa e Nandamudi Lankalapalli Vijaykumar,
aprovada em 26 de maio de 2006.

INPE
São José dos Campos
2007

Publicado por:

esta página é responsabilidade do SID

Instituto Nacional de Pesquisas Espaciais (INPE)

Gabinete do Diretor – (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 – CEP 12.245-970

São José dos Campos – SP – Brasil

Tel.: (012) 3945-6911

Fax: (012) 3945-6919

E-mail: pubtc@sid.inpe.br

**Solicita-se intercâmbio
We ask for exchange**

Publicação Externa – É permitida sua reprodução para interessados.



MINISTÉRIO DA CIÊNCIA E TECNOLOGIA
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

INPE-14670-TDI/1223

**ESTUDO DA VARIABILIDADE DE SÉRIES TEMPORAIS
AMBIENTAIS ATRAVÉS DE ESPECTROS DE ASSIMETRIA
IMPLEMENTADOS EM SOFTWARE CIENTÍFICO LIVRE**

Eduardo Batista de Moraes Barbosa

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada,
orientada pelos Drs. Reinaldo Roberto Rosa e Nandamudi Lankalapalli Vijaykumar,
aprovada em 26 de maio de 2006.

INPE
São José dos Campos
2007

681.3.019

Barbosa, E. B. M.

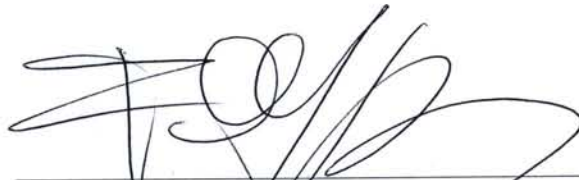
Estudo da variabilidade de séries temporais ambientais através de espectros de assimetria implementados em software científico livre / Eduardo Batista de Moraes Barbosa. – São José dos Campos: INPE, 2006.

107p. – (INPE-14670-TDI/1223)

1.Análise de séries temporais. 2.Transformada de ondeletas. 3.Processamento de dados. 4.Código aberto. 5.Programação de computador. I.Título.

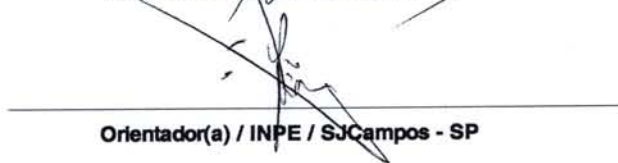
Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de **Mestre** em
Computação Aplicada

Dr. Fernando Manuel Ramos



Presidente / INPE / SJC Campos - SP

Dr. Reinaldo Roberto Rosa



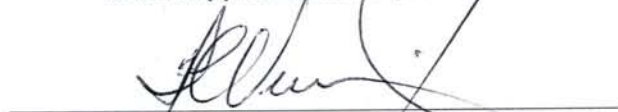
Orientador(a) / INPE / SJC Campos - SP

Dr. Nandamudi Lankalapalli Vijaykumar



Orientador(a) / INPE / SJC Campos - SP

Dr. Haroldo Fraga de Campos Velho



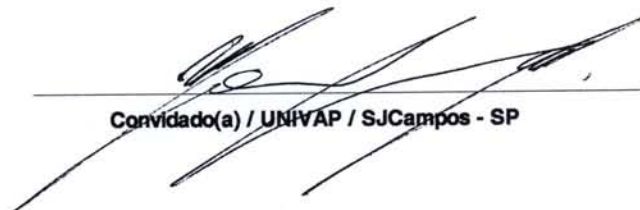
Membro da Banca / INPE / SJC Campos - SP

Dr. Edvaldo Nogueira Júnior



Convidado(a) / UFBA / Salvador - BA

Dr. Maurício José Alves Bolzan



Convidado(a) / UNIVAP / SJC Campos - SP

Aluno (a): **Eduardo Batista de Moraes Barbosa**

São José dos Campos, 26 de maio de 2006

AGRADECIMENTOS

Gostaria de expressar meu agradecimento primeiramente a DEUS por me guiar sempre, dando-me a orientação necessária para superar todas as dificuldades que aparecem em meu caminho. "DEUS dirige todos os acontecimentos, porque está em tudo".

Expresso meu carinho e gratidão às seguintes pessoas:

Meus pais, minha irmã e minha namorada pelo incentivo e compreensão desde o início até a conclusão desta jornada;

Meus orientadores Prof. Dr. Reinaldo e Vijaykumar pelo apoio na definição do trabalho, pela confiança em mim depositada, pelo incentivo e, principalmente, pela orientação;

Dr. Maurício Bolzan, Dr. Arcilan Assireu, Dr. Márcio Vianna e Dr. Ademilsson Zanandrea pelas explicações, pelos dados e pelo apoio;

Dr. Javier Tomasella pela paciência;

Professores pelo comprometimento na busca da qualidade de trabalho, conceito e conhecimento;

Amigos pela ajuda durante o curso;

Banca examinadora pelas sugestões e comentários para a melhora desta dissertação.

RESUMO

A análise de séries temporais (ST) curtas ($\sim 10^3$ pontos) e não-estacionárias, em geral, comuns na natureza, exige a utilização de abordagens robustas da matemática. Duas novas técnicas estatísticas computacionais que merecem destaque são a transformada em ondeletas e a análise de padrões-gradiente. O presente trabalho tem por objetivo investigar as características dessas técnicas implementadas em software científico de domínio público (SCDP) *Scilab*. As técnicas foram aplicadas em diferentes ST e seus resultados mostraram-se coerentes, quando comparados com os resultados das implementações originais, desenvolvidas em ambientes distintos de programação MATLAB® e *Interactive Data Language* (IDL). Realiza-se uma discussão sobre a importância do desenvolvimento e utilização de SCDP no contexto dos cursos de pós-graduação e projetos de pesquisa realizados no Instituto Nacional de Pesquisas Espaciais (INPE). Adicionalmente, propõe-se uma metodologia alternativa para a análise de ST, apropriada e robusta para ST curtas. Esta metodologia, que complementa a caracterização de diferentes padrões de variabilidades em ST curtas, consiste na utilização conjunta da transformada de ondeletas discretas e da análise de padrões-gradiente, para obtenção do espectro de assimetrias. Resultados preliminares, obtidos com a aplicação da metodologia em séries de ciências atmosféricas, conseguiram caracterizar, de maneira quantitativa, diferentes padrões de variabilidades associados aos seguintes regimes não-lineares: caótico, reativo-difusivo e turbulento.

THE STUDY OF ENVIRONMENT TIME SERIES VARIABILITY THROUGH ASYMMETRY SPECTRA IMPLEMENTED USING FREE SCIENTIFIC SOFTWARE

ABSTRACT

The analysis of short ($\sim 10^3$ points) and non-stationary time series (TS), quite common in the nature, demands the use of robust mathematical techniques. Two new computational statistical techniques that deserve attention are wavelet transforms and the gradient pattern analysis. The objective of this work is to investigate the characteristics of these techniques implemented within a free scientific software environment *Scilab*. The techniques were applied in different TS and the results obtained from this environment were similar when compared with the results obtained from the same techniques implemented in proprietary software environments MATLAB® and Interactive Data Language (IDL). There has been discussions on the importance of the development and use of free scientific software in the context of the under-graduate courses and carried out through research projects at the National Institute for Space Research (INPE). Additionally, an alternative methodology is considered for the analysis of TS, very appropriate and robust for short TS. This methodology, that complements the characterization of different standards of variabilities in short TS, consists of the joint use of the discrete wavelet transforms and the gradient pattern analysis, for obtaining the asymmetry spectra. Preliminary results from the application of this methodology in series from the fields of atmospheric sciences, characterized in a quantitative way, different standards of variability associates with the following nonlinear regimes: chaotic, reactive-diffusive and turbulent.

SUMÁRIO

Pág.

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE SIGLAS

CAPÍTULO 1 – INTRODUÇÃO	21
1.1 Estudo de Séries Temporais	21
1.2 Softwares Científicos.....	24
1.3 Objetivos.....	26
1.4 Organização da Dissertação.....	27
CAPÍTULO 2 – SÉRIES TEMPORAIS	29
2.1 Técnicas Convencionais para Análise de Séries Temporais	31
2.1.1 Função de Autocorrelação.....	31
2.1.2 Análise Espectral	32
CAPÍTULO 3 – APRESENTAÇÃO DOS CONJUNTOS DE DADOS	35
3.1 Séries Temporais Canônicas.....	35
3.1.1 Modelo Gaussiano	35
3.1.2 Padrão de Variabilidade Caótico 1d (Mapa Logístico)	36
3.1.3 Padrão de Variabilidade Caótico 2d (Mapa de Hénon).....	38
3.1.4 Padrão de Variabilidade Caótico 3d (Modelo de Lorenz).....	39
3.1.5 Padrão de Variabilidade Reativo-Difusivo (Equação de Kuramoto-Sivashinsky)	41
3.1.6 Padrão de Variabilidade Turbulento (Dados de Velocidade do Vento – Projeto Lba).....	42
3.2 Séries Temporais Ambientais.....	44
3.2.1 Sistema Integrado de Monitoramento Ambiental.....	44
3.2.2 Dados Oceanográficos.....	48
CAPÍTULO 4 – TÉCNICAS ESTATÍSTICAS EM SCDP <i>SCILAB</i>	51
4.1 Softwares de Domínio Público.....	51

4.2 Software Científico de Domínio Público: <i>Scilab</i>	53
4.2.1 Descrição do Funcionamento do <i>Scilab</i>	54
4.3 Resultados Preliminares	56
4.3.1 Estatística Convencional	56
4.3.1.1 Média, Variância e Desvio-Padrão	56
4.3.1.2 Autocorrelação	57
4.3.2 Implementação de Técnicas Estatísticas Avançadas em <i>Scilab</i>	58
4.3.2.1 Transformada de Ondeletas em <i>Scilab</i>	58
4.3.2.2 Análise de Padrões-Gradiente em <i>Scilab</i>	61
CAPÍTULO 5 – APLICAÇÃO CONJUNTA DA TRANSFORMADA DE ONDELETAS E ANÁLISE DE PADRÕES-GRADIENTE: ESPECTRO DE ASSIMETRIAS	65
5.1 Decomposição Tempo-Frequência por Transformada de Ondeletas	65
5.2 Cálculo das Assimetrias das Componentes Espectrais.....	68
5.3 Espectro de Assimetrias	69
5.3.1 Primeira Abordagem	70
5.3.2 Segunda Abordagem	76
5.4 Resultados Preliminares – Primeira e Segunda Abordagens.....	79
CAPÍTULO 6 – CONCLUSÃO	81
6.1 Trabalhos Futuros	82
REFERÊNCIAS BIBLIOGRÁFICAS	83
APÊNDICE A – TEORIA DE ONDELETAS	91
APÊNDICE B – ANÁLISE DE PADRÕES-GRADIENTE	99
B.1 Análise de Padrões-Gradiente de Séries Temporais (Gpa1d).....	102
APÊNDICE C – SCILAB NA INTERNET	105
C.1 Algoritmos Desenvolvidos em <i>Scilab</i>	107

LISTA DE FIGURAS

	<u>Pág.</u>
3.1 – Série temporal curta gaussiana (1024 pontos). No detalhe, uma sub-amostra com 128 pontos	36
3.2 – Série temporal curta (1024 pontos) com padrão de variabilidade caótico 1D (Mapa logístico). No detalhe, uma sub-amostra com 128 pontos.....	37
3.3 – Série temporal curta (1024 pontos) com padrão de variabilidade caótico 2D (Mapa de Hénon). No detalhe, uma sub-amostra com 128 pontos.....	39
3.4 – Série temporal curta (1024 pontos) com padrão de variabilidade caótico 3D (Modelo de Lorenz). No detalhe, uma sub-amostra com 128 pontos.....	40
3.5 – Série temporal curta (1024 pontos) com padrão de variabilidade reativo-difusivo (Equação de Kuramoto-Sivashinsky). No detalhe, uma sub-amostra com 128 pontos.	42
3.6 – Anemômetro Sônico tridimensional instalado na torre micrometeorológica da Reserva Biológica do Jaru. (FONTE: Bolzan, 2002)	43
3.7 – Série temporal curta com padrão de variabilidade turbulento (Velocidade do Vento – projeto LBA) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.....	44
3.8 – Fotografia de uma PCD instalada no reservatório de Serra da Mesa. (FONTE: Stech et al., 2004)	45
3.9 – Localização dos reservatórios selecionados para a realização do projeto SIMA..	46
3.10 – Série temporal curta de temperatura do ar (reservatório de Manso) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.....	47
3.11 – Série temporal curta de temperatura do ar (reservatório de Serra da Mesa) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.....	47
3.12 – Série temporal curta de amplitudes de correntes oceânicas (Componente Leste) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.....	48
3.13 – Série temporal curta de amplitudes de correntes oceânicas (Componente Norte) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.....	49
4.1 – Janela inicial do Scilab 3.1.1, versão pré-compilada para Windows 2000	53
4.2 – Scipad editor de programas fornecido em conjunto com o Scilab	55
4.3 – Diagramas de dispersão, em Scilab, para dois conjuntos de dados atmosféricos. (a) Temperatura do ar (Serra da Mesa); (b) Amplitudes de correntes oceânicas (componente Leste)	57
4.4 – Resultados da transformada de ondeletas aplicada em dados de temperatura do ar no reservatório de Serra da Mesa. (a) Scilab; (b) MATLAB®	60

4.5 – Resultados da transformada de ondeletas aplicada em dados de amplitudes de correntes oceânicas (componente leste). (a) Scilab; (b) MATLAB®	61
5.1 – Exemplos de decomposições da série gaussiana (topo) com a ondeleta discreta de Haar	66
5.2 – Exemplos de decomposições da série gaussiana (topo) com a ondeleta discreta de Daubechies	67
5.3 – Padrões de simétricos elementares (9 pontos) e suas respectivas matrizes e campos-gradiente.....	68
5.4 – Padrões de assimétricos elementares (9 pontos) e suas respectivas matrizes e campos-gradiente.....	69
5.5 – Algoritmo generalizado para obtenção do espectro de assimetrias (Primeira abordagem).....	70
5.6 – (a) Série original; (b) Somatória das cinco últimas decomposições do sinal	71
5.7 – Espectro de assimetrias para a série gaussiana.....	74
5.8 – Espectro de assimetria para dados canônicos com padrão de variabilidade caótico 2D (Mapa de Hénon).....	75
5.9 – Algoritmo generalizado para obtenção do espectro de assimetrias (Segunda abordagem).....	76
5.10 – Espectro de assimetrias para a série gaussiana.....	78
5.11 – Espectro de assimetria para dados canônicos com padrão de variabilidade caótico 2D (Mapa de Hénon).....	78
A.1 – Variações proporcionais dos intervalos de tempo e de frequência. (FONTE: Graps, 1995)	93
A.2 – Função ondeleta “Haar”. (FONTE: Graps, 1995)	94
A.3 – Função ondeleta “Daubechies”. (FONTE: Graps, 1995)	94
A.4 – Função ondeleta “Morlet”. Parte real: linha cheia. Parte imaginária: linha pontilhada. (FONTE: Torrence e Compo, 1998).....	95
A.5 – Função ondeleta “Chapéu Mexicano”. Parte real: linha cheia. Parte imaginária: linha pontilhada. (FONTE: Torrence e Compo, 1998).....	96
B.1 – Momentos Gradientes: (a) Momento escalar, (b) Momento gradiente de primeira ordem, (c) Momentos gradiente de diversidades de módulo e fase e (d) Momento da medida global da norma e da fase.....	99
B.2 – Matrizes representativas de campos vetoriais assimétricos. (a) 6 vetores assimétricos, (b) e (c) todos os vetores são assimétricos.....	101
B.3 – Campo de triangulação para (b) e (c) da Figura B.2.....	102

B.4 – (a) Típico padrão de variabilidade, (b) Matriz correspondente, (c) Campo gradiente correspondente (9 vetores assimétricos), (d) Campo de triangulação com 20 linhas conectoras.....	103
C.1 – Página principal do sítio do Scilab na Internet. Destaque para a seção <i>Downloads</i> à esquerda da página.....	104
C.2 – Seção <i>Downloads</i> encontrada no sítio do Scilab. Destaque para a versão desenvolvida para o sistema operacional Windows	105
C.3 – Opções de <i>Downloads</i> da versão 4.0 do Scilab, recomendada para instalações pelo sítio do software na Internet	105

LISTA DE TABELAS

	<u>Pág.</u>
4.1 – Parâmetros estatísticos calculados para dois conjuntos de dados atmosféricos em Scilab	56
4.2 – Autocorrelações de diferentes conjuntos de dados atmosféricos	58
4.3 – Resultados da GPA1D para o padrão de variabilidade caótico 3D (modelo de Lorenz)	63
4.4 – Resultados da GPA1D para o padrão de variabilidade turbulento (velocidade do vento – projeto LBA).....	63
5.1 – Resultados de $\langle g_1^a \rangle$ locais, obtidos após a aplicação da GPA1D nas séries de decomposições.....	72
5.2 – Resultados de g_1^a globais, obtidos após a aplicação da GPA1D nas séries de decomposições.....	72
5.3 – Extensão normalizada da assimetria ($g_1^{a,R}$), calculada para a série gaussiana.....	73
5.4 – Extensão normalizada da assimetria ($g_1^{a,R}$) calculadas para a ST canônica com padrão de variabilidade caótico 2D (Mapa de Hénon).....	74
5.5 – Distâncias de Kullback-Leibler entre os espectros de assimetria (série gaussiana e dados canônicos)	76
5.6 – Resultados de $\langle g_1^a \rangle$ locais e g_1^a globais, obtidos após a aplicação da GPA1D nas séries de decomposições	77
5.7 – Distâncias de Kullback-Leibler entre os espectros de assimetria (série gaussiana e dados canônicos)	79
5.8 – Distâncias de Kullback-Leibler entre os espectros de assimetria (séries gaussiana e estudos de casos) – Primeira abordagem.....	79
5.9 – Distâncias de Kullback-Leibler entre os espectros de assimetria (séries gaussiana e estudos de casos) – Segunda abordagem.....	79

LISTA DE SIGLAS

DFA	<i>Análise das Flutuações Destendenciadas</i>
ENPC	<i>École Nationale des Ponts et Chaussées</i>
FSF	<i>Free Software Foundation</i>
FT	<i>Transformada de Fourier</i>
GPL	<i>GNU Public License</i>
GNU	<i>Gnu's Not Unix</i>
GPA	<i>Gradient Pattern Analysis</i>
GPAID	<i>GPA para Análises de Séries Temporais</i>
GWS	<i>Global Wavelet Spectrum</i>
IDL	<i>Interactive Data Language</i>
INPE	<i>Instituto Nacional de Pesquisas Espaciais</i>
INRIA	<i>Institut de Recherche en Informatique et en Automatique</i>
LAC	<i>Laboratório Associado de Computação e Matemática Aplicada</i>
LBA	<i>Large Scale Biosphere-Atmosphere Experiment in Amazonia</i>
PCD	<i>Plataforma de Coleta de Dados</i>
SCDP	<i>Software Científico de Domínio Público</i>
SIMA	<i>Sistema de Monitoramento Ambiental</i>
ST	<i>Série Temporal</i>
UNIVAP	<i>Universidade do Vale do Paraíba</i>

CAPÍTULO 1

INTRODUÇÃO

1.1 Estudo de Séries Temporais

Grande parte das medidas científicas provenientes da observação e/ou simulação é registrada através de séries temporais. Nas últimas três décadas, séries temporais passaram a ser analisadas dentro de um novo paradigma, isto é, com o desenvolvimento de novos materiais, instrumentos de medidas e dispositivos computacionais avançados, tornou-se possível obter estas séries em alta resolução e alta sensibilidade, para auxiliar na identificação de detalhes não considerados em análises anteriores. Devido à altas resoluções (temporal, espectral e espacial) das observações e dos experimentos numéricos, os cientistas passaram a analisar e modelar a *variabilidade complexa* que constitui toda a amostra ou sub-amostras das amplitudes que compõem a série. É usual interpretar como variabilidade complexa regimes de variação que compreendem persistência, anti-persistência e intermitências no domínio temporal. Portanto, nas últimas três décadas novas ferramentas foram desenvolvidas e introduzidas na literatura com a finalidade de auxiliar a caracterização de padrões não-lineares, possivelmente, associados à natureza da fonte responsável pela produção da variabilidade complexa observada.

Um exemplo clássico de dados que apresentam variabilidade complexa é aquele relacionado a observações de variáveis atmosféricas. Em geral, um fenômeno geofísico natural envolve diversas variáveis correlacionadas entre si (e também auto-correlacionadas) por meio de leis estocásticas e/ou determinísticas que podem envolver processos não-lineares. De acordo com resultados recentes publicados na literatura (Emery e Thomson, 2001; Assireu et al., 2002; Ramos et al., 2004), essa variabilidade complexa impõe fortes restrições à qualidade das previsões, que podem ser importantes,

por exemplo, na caracterização de processos atmosféricos. Além disso, dados atmosféricos possuem características particulares com relação ao formato de organização, isto é, podem ser representados por pontos de grades geo-referenciados, pontos de estação, séries temporais, séries espaciais, etc.

O presente trabalho está focado na análise de séries temporais relacionadas a ciências atmosféricas. Tais séries são, usualmente, não-estacionárias e apresentam variabilidades complexas, ou seja, nestas séries os momentos estatísticos, em geral, variam para a maioria dos segmentos tomados como sub-amostras. Por isso, a sua análise exige abordagens robustas da matemática, para possibilitar a compreensão dos principais processos físicos relacionados ao sistema subjacente.

Através da literatura, pode-se observar que a utilização de novas técnicas estatísticas, continuamente investigadas e desenvolvidas nos últimos anos, tem introduzido novos paradigmas à teoria da análise de séries temporais com variabilidade complexa (Assireu, et al., 2002). As principais metodologias envolvem técnicas computacionalmente custosas e de difíceis implementações, que dependem de um número mínimo de medidas seqüenciais (aproximadamente 10^4 pontos). Essas restrições comprometem a análise de séries curtas (aproximadamente $10^2 - 10^3$ pontos), uma outra particularidade das variáveis atmosféricas.

Como exemplos podem ser citadas as metodologias que fornecem parâmetros que convergem dependendo do cálculo de desvios em relação a um dado valor de referência (a média, por exemplo). As principais são as Leis de Potência obtidas a partir de Funções de Densidade de Probabilidade (Osborne e Provenzale, 1989) e as Dimensões de Correlação obtidas a partir de Reconstruções do Espaço de Fases (Schreiber, 1998).

Várias metodologias são utilizadas com o objetivo de caracterizar, de maneira precisa e robusta, a existência e a intensidade da autocorrelação do sinal temporal (curto, médio e de longo alcance) explicitando por isso a riqueza de escalas (temporais e espaciais) que podem estar envolvidas nos processos subjacentes: intermitência e turbulência (Frish,

1995), auto-organização fora do equilíbrio (Bak et al., 1988), difusões anômalas (Swinney e Tsallis, 2004), caoticidade (Peitgen et al., 1992) e regimes reativo-difusivos (Cross e Hohenberg, 1993).

Um exemplo de metodologia recente, que tem levado a um número crescente de aplicações, é a Análise das Flutuações Destendenciadas (DFA) (*Detrended Fluctuation Analysis*) (Peng et al., 1992). Técnicas como esta tem como objetivo o cálculo de uma flutuação estatística $F(n)$, onde n representa o tamanho de uma janela para mapear o conjunto de medidas. Variando o tamanho de n , as flutuações podem ser caracterizadas através de um expoente de escala obtido a partir da curva ajustada ao gráfico $\log F(n)$ versus $\log n$. Para obtenção de resultados robustos é necessário variar o tamanho da janela envolvendo desde um número pequeno até um número grande de pontos cobrindo todo conjunto que representa a série temporal. Dessa maneira, quando as séries são extensas (aproximadamente 10^4 pontos), é possível, por meio do expoente de escala, caracterizar a série, de forma mais robusta, através de suas correlações de longo alcance.

Recentemente, a abordagem multi-fractal da DFA, tratada originalmente por metodologias envolvendo transformada de ondeletas, proporcionou uma generalização capaz de identificar a fonte da multi-fractalidade quando esta está presente no sinal (Kantelhardt et al., 2002). Entretanto, todas essas metodologias perdem robustez quando aplicadas a séries temporais curtas (aquelas constituídas por uma quantidade de pontos que compromete a universalidade dos valores para os expoentes de escala encontrados).

Como o objetivo principal desta dissertação é introduzir uma metodologia alternativa para análise de séries temporais curtas, com exceção das transformadas de ondeletas, as demais técnicas supracitadas não serão abordadas.

Para introduzir a metodologia alternativa, considerando a variabilidade complexa de séries temporais curtas, duas técnicas a serem utilizadas correspondem à Transformada de Ondeletas (*Wavelets Transform*) (Farge, 1992) e à Análise de Padrões-Gradiente

(*Gradient Pattern Analysis*) (Rosa et al., 1999), adaptada para séries restritas ao domínio temporal (Assireu et al., 2002).

Nas últimas décadas a transformada de ondeletas tornou-se uma ferramenta comum para entender as variações locais em uma série temporal. Sua idéia central, no contexto de análises de sinais, consiste em decompor uma série temporal em diferentes níveis de resolução tempo-frequência e, então, determinar as componentes de variabilidade dominantes (Percival e Walden, 2002). Na literatura atual pode-se encontrar um extenso conjunto de trabalhos relacionados com as transformadas de ondeletas. Uma abordagem teórica atualizada sobre a técnica é apresentada por Graps (Graps, 1995) destacando suas potencialidades de aplicações em diversas áreas do conhecimento como, por exemplo, processamento de imagens, compressão e filtragem de dados e caracterização de séries temporais. Neste artigo, destaca-se a vantagem da transformada de ondeletas, sobre os métodos tradicionais de Fourier, quando aplicada na análise de dados de ciências atmosféricas que, em geral, apresentam complicadas descontinuidades e pulsos ao longo do tempo.

A análise de padrões-gradiente (GPA), formalmente introduzida por Rosa et al. (1999) e adaptada por Assireu et al. (2002) para análise de séries temporais (GPA1D), adota o pressuposto de que a complexidade de uma série temporal é caracterizada pelo seu grau de fragmentação assimétrica (Assireu et al., 2004). Ou seja, à medida que o padrão de variabilidade de uma série temporal se torna mais irregular (no domínio da amplitude e/ou da frequência) o valor de assimetria obtido através do primeiro momento gradiente tende a aumentar, resultando em uma das conjecturas relacionadas a essa metodologia.

1.2 Softwares Científicos

Em instituições de ensino e pesquisa é comum a utilização de produtos de softwares proprietários como, por exemplo, MATLAB®, Mathematica e *Interactive Data Language* (IDL), para a aquisição, o processamento, a visualização e a análise de dados científicos. Em geral, estes produtos são adotados devido aos bons índices de sucesso

tecnológico e alto desempenho, obtidos em suas áreas de atuação. Embora atraentes, sob ponto de vista de software, e desenvolvidos cuidadosamente para mercados específicos, os produtos de softwares proprietários raramente alcançam níveis elevados de satisfação.

Nos últimos anos, produtos de softwares proprietários, destinados à computação científica, começaram a se deparar com um novo desafio, isto é, a concorrência com produtos desenvolvidos sob conceito de softwares de domínio público. Em linhas gerais, o conceito de softwares públicos consiste em garantir aos possíveis usuários, liberdade para acesso efetivo ao software, enquanto que os produtos proprietários, amparados por leis específicas, visam restringir esse acesso por meio da comercialização de licenças com valores diferenciados (demonstrativas, estudantis, multi-usuários, super-licenças, etc).

Além de gratuitos, característica dos produtos de softwares de domínio público relevante para as instituições públicas e privadas de ensino e pesquisa, outra característica que pode ser considerada em sua adoção corresponde à flexibilidade para a disseminação e a utilização dos produtos. Conseqüentemente, todas as burocracias financeiras são sumariadas, uma vez que não existem necessidades de compra e licenciamento, e a propriedade intelectual, bem como, o conhecimento podem ser compartilhados.

Uma característica relevante em grande parte dos produtos de softwares de domínio público, principalmente, na área de computação científica, consiste na capacidade de agregação aos produtos de softwares proprietários. Neste contexto, existem diferentes alternativas de boa qualidade como, por exemplo, o Octave e o *Scilab*. Em particular, o ambiente de desenvolvimento *Scilab*, oferece um ambiente integrado de computação científica, pode, em tese, reproduzir as principais funções e aplicações do MATLAB® e, também, do IDL.

1.3 Objetivos

O presente trabalho tem por objetivo implementar, em *software científico de domínio público* (SCDP), as técnicas de transformada de ondeletas e de GPA1D. Estas técnicas serão aplicadas na análise de séries temporais (ST) curtas e não-estacionárias associadas a dados de ciências atmosféricas (meteorologia, oceanografia, etc.). Cabe ressaltar que ambas as técnicas foram desenvolvidas e divulgadas, originalmente, no contexto do software científico proprietário, MATLAB® e IDL, respectivamente. Dessa maneira, atualmente, a sua utilização compreende a aquisição de licenças comercializadas pelos representantes legais proprietários das marcas, que são pessoas jurídicas com fins lucrativos.

Além de implementar as técnicas de transformada de ondeletas de GPA1D em SCDP, será introduzida uma metodologia alternativa para a análise de ST, apropriada e robusta para ST curtas, que consiste na utilização conjunta da transformada de ondeletas e da GPA1D, para obtenção do *espectro de assimetrias*. A metodologia é baseada nas decomposições de uma série através da ondeleta discreta de Daubechies (Daubechies, 1988) e na aplicação da GPA1D para obtenção das assimetrias das séries de decomposições.

Em resumo, este trabalho compreenderá: (i) o estudo e a implementação da transformada de ondeletas e da GPA1D em ambiente de SCDP *Scilab*; (ii) a geração de ST canônicas* para validação das técnicas; (iii) a aplicação das técnicas, através do espectro de assimetrias, em estudos de casos envolvendo dados experimentais relacionados aos projetos ambientais desenvolvidos no Instituto Nacional de Pesquisas Espaciais (INPE) e (iv) a introdução de uma discussão sobre a importância da utilização de SCDP no contexto dos cursos de Pós-Graduação e projetos de pesquisa realizados no INPE.

* Séries canônicas: Sub-amostras normalizadas, com 1024 pontos, representando padrões de variabilidade conhecidos, isto é, periódico, caótico, turbulento e reativo-difusivo.

1.4 Organização da Dissertação

A presente dissertação de mestrado está organizada da seguinte maneira.

No Capítulo 2, a partir de uma breve introdução ao conceito de ST é apresentada, de acordo com a literatura, uma relação das principais técnicas para análise de ST.

No Capítulo 3 são apresentados os conjuntos de dados, canônicos e referentes aos estudos de caso, utilizados neste trabalho.

As técnicas de transformada de ondeletas e de GPA1D são apresentadas, de forma detalhada, no Capítulo 4, bem como, a sua implementação e validação em SCDP *Scilab*.

No Capítulo 5 é introduzida a metodologia para obtenção do espectro de assimetrias, que pode auxiliar na caracterização de diferentes padrões de variabilidades representados pelas ST, inclusive os estudos de caso, apresentadas no Capítulo 3.

Finalmente, o Capítulo 6 conclui o texto da dissertação apresentando um sumário dos principais resultados, conclusões e futuras direções a serem exploradas.

CAPÍTULO 2

SÉRIES TEMPORAIS

Uma ST pode ser definida como um conjunto de observações em função do tempo (Morettin e Toloi, 2004), isto é,

$$\{X_t\}, t = 1, 2, 3, \dots, N;$$

onde t representa a variação temporal da ST e N , o comprimento desta.

Dados de ST podem surgir de diferentes áreas de conhecimento como, por exemplo,

- Meteorologia e Oceanografia: precipitação, temperatura, velocidade do vento, velocidade e amplitudes de correntes marítimas, etc;
- Astrofísica: curvas de luz de objetos estelares, número de manchas solares, explosões solares observadas em radiofrequência e altas energias, etc;
- Geofísica espacial: pulsações geomagnéticas, cintilações ionosféricas, etc.;
- Economia: índice da Bolsa de Valores, dinâmica de derivativos, etc.;
- Medicina: eletrocardiograma, eletro-encefalograma, etc.

Uma das características mais relevantes nestes tipos de dados é a dependência entre as observações vizinhas e, em geral, o interesse concentra-se na análise e modelagem desta dependência fenomenológica (Moretin e Toloi, 2004) ou ainda, na identificação de sistemas matemáticos (Aguirre, 2004).

Os objetivos de se realizar análises em ST são variados. Em algumas situações, pode ser, simplesmente, obter previsões sobre valores ou tendências futuras e em outras, uma

investigação da relação entre estruturas de duas ou mais séries. Estes objetivos, em geral, podem ser sintetizados da seguinte maneira:

- Caracterização: determinar as propriedades da série (padrões de tendência, existência de variação sazonal, regimes não-lineares, autocorrelação, leis de potência, etc.);
- Previsão: prever variações passadas e futuras em relação a uma seqüência de medidas;
- Modelagem: aplicar ferramentas para modelagem matemática e/ou fenomenológica da série.
- Controle de processos: controle estatístico de processos e qualidade dos dados;
- Comparações: usar a variação numa série para explicar a variação em outra.

Séries que apresentam comportamentos periódico ou quase-periódico podem ser caracterizadas e previstas através de funções trigonométricas. Estes tipos de ST são classificadas como determinísticas, uma vez que seus valores futuros podem ser exatamente representados por uma função matemática. Para séries mais complexas, com comportamento irregular e/ou aleatório, tanto a caracterização como a previsão, ficam restritas às estimativas de parâmetros estatísticos. Em geral, os padrões de variabilidades destas séries são classificados como aleatórios ou estocásticos.

A análise de ST complexas e não-estacionárias exige abordagens robustas da matemática, com a finalidade de compreender os processos, geralmente físicos, que atuam no sistema subjacente. Parâmetros estatísticos de ST podem ser utilizados para determinar similaridades e diferenças entre conjuntos de dados distintos. Entretanto, para séries curtas (aproximadamente 10^3 pontos) e não-estacionárias como, por exemplo, geradas através de modelos matemáticos e/ou relacionadas à observação de variáveis atmosféricas, as medidas estatísticas convencionais são pouco representativas (Assireu et al., 2002).

2.1 Técnicas Convencionais para Análise de Séries Temporais

Dentro de uma abordagem clássica para a análise e tratamento de ST podem-se encontrar diferentes metodologias, cujas mais clássicas são as análises da função de autocorrelação e espectral.

2.1.1 Função de Autocorrelação

A função de autocorrelação de uma ST $\{X_t\}$ pode ser definida como,

$$\rho = \frac{\sum_{t=0}^{N-k-1} (X_t - \bar{x})(X_{t+k} - \bar{x})}{\sum_{t=0}^{N-1} (X_t - \bar{x})^2} \quad (2.1)$$

, onde N representa o comprimento da ST e \bar{x} é o valor esperado (2.2) das observações, calculada para a variação temporal (*lag*) k . O coeficiente de autocorrelação (ρ) de uma ST varia entre -1 e 1, inclusive. Se ρ assume o valor 1, então, diz-se que as duas variáveis medidas possuem uma autocorrelação positiva absoluta, caso contrário, se $\rho = -1$, diz-se que a autocorrelação negativa é absoluta. Quando ρ assume valor 0, não existe autocorrelação entre as variáveis, ou seja, autocorrelação nula.

O valor esperado de uma ST (média aritmética) é definido como um valor típico ou representativo dos dados. Isto é, através da média aritmética pode-se apresentar o valor do ponto em torno do qual os dados se distribuem.

$$\bar{x} = \frac{1}{N} \sum_{t=0}^{N-1} X_t, \quad (2.2)$$

O grau de dispersão (variância) dos dados pode ser medido pelos desvios (dos dados) em relação ao valor esperado, considerando-se o tamanho da amostra. A variância (σ^2) de uma ST é definida como:

$$\sigma^2 = \frac{1}{N-1} \sum_{t=0}^{N-1} (X_t - \bar{x})^2 \quad (2.3)$$

, onde $(N-1)$ representa o grau de liberdade dos dados e \bar{x} é o valor esperado da série. Entretanto, como medida de dispersão, a variância tem a desvantagem de apresentar a unidade de medida igual ao quadrado da unidade de medida dos dados. Por isso, utiliza-se o desvio padrão (σ), que apresenta a mesma unidade de medida dos dados, e é definido como a raiz quadrada da variância, isto é,

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{t=0}^{N-1} (X_t - \bar{x})^2}. \quad (2.4)$$

2.1.2 Análise Espectral

A análise de Fourier ou análise espectral tem sido utilizada, tradicionalmente, na resolução de equações diferenciais parciais que aparecem na física-matemática, como equação do calor e de ondas.

No século XIX, Joseph Fourier, com sua teoria de análise de frequências, afirmou que qualquer função periódica $f(x)$ pode ser expressa pela somatória de senos e co-senos, ou seja,

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx)). \quad (2.5)$$

Os coeficientes a_0 , a_k e b_k são calculados, respectivamente, por,

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx, \quad (2.6)$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx) dx, \quad (2.7)$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx) dx. \quad (2.8)$$

Portanto, representações em séries de Fourier, de funções ou de dados, como superposição das funções seno e co-seno têm sido muito úteis para soluções analítica e numérica para análises e tratamento de sinais.

A transformada de Fourier (FT) é um método para decomposição de uma ST em domínio de tempo em suas componentes de frequência. As equações que representam a FT podem ser escritas como segue,

$$f(t) = \int_{-\infty}^{+\infty} F(\omega) e^{j\omega t} d\omega \quad (2.9)$$

, onde

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-j\omega t} dt \quad (2.10)$$

, onde t representa o tempo, ω é a frequência em radianos, $f(t)$ é um sinal de entrada contínuo e $F(\omega)$ é o espectro de Fourier validado sobre as funções base seno e co-seno, como indicado por $e^{j\omega}$. Nas equações acima, as funções trigonométricas foram expressas em sua forma complexa.

Todo sinal em domínio de tempo, não importando o quanto complexo, pode ser separado em componentes senoidais. Como resultado, obtém-se um conjunto de valores $F(\omega)$ indicando as componentes de frequência e sua respectiva amplitude. Os resultados da FT não fornecem informações explícitas sobre a duração, início ou final das componentes de frequência. Nesse sentido, a FT fornece uma imagem global da frequência do sinal contido no tempo. Portanto, pode-se observar uma deficiência da FT se aplicada em ST não-estacionárias, o que é comum em dados de ciências atmosféricas.

Diante desse contexto, observa-se a necessidade de investigar outras técnicas que viabilizem a realização de análises eficazes tanto em dados de ST estacionárias como não-estacionárias. Diversas técnicas têm sido propostas para caracterizar propriedades de diferentes séries de natureza não-estacionária. Para uma relação detalhada ver, por exemplo, Campanharo (2005).

CAPÍTULO 3

APRESENTAÇÃO DOS CONJUNTOS DE DADOS

Durante a realização deste trabalho foram utilizadas diversas ST obtidas através (i) de experimentos computacionais e observacionais, representando diferentes padrões canônicos de variabilidades temporais (multi-periódico, turbulento, caótico e reativo-difusivo) e (ii) de estudos de casos, associados a variáveis atmosféricas, obtidas a partir de projetos vinculados ao INPE.

3.1 Séries Temporais Canônicas

Nas próximas seções serão apresentadas as definições das ST (i) multi-periódica (padrão de referência), (ii) logística, (iii) de Hénon, (iv) de Lorenz e (v) de Kuramoto-Sivashinsky. Adicionalmente, será apresentada uma série de velocidade do vento obtida através do projeto *Large Scale Biosphere-Atmosphere Experiment in Amazonia* (LBA), utilizada neste trabalho para representar dados com padrão de variabilidade do tipo turbulento. Estes conjuntos de dados canônicos, constituídos por 10^4 pontos normalizados entre 0 e 1, formam um arquivo de ST implementado durante a realização desse trabalho, com a finalidade de validar a proposta apresentada.

3.1.1 Modelo Gaussiano

No âmbito deste trabalho, o grau de irregularidade associado ao padrão de variabilidade de uma ST, foi definido através de um padrão gaussiano (média zero). Na Figura 3.1 é apresentada a série curta completa (1024 pontos) e, no detalhe, uma aproximação correspondente à sub-amostra da mesma constituída por 128 pontos.

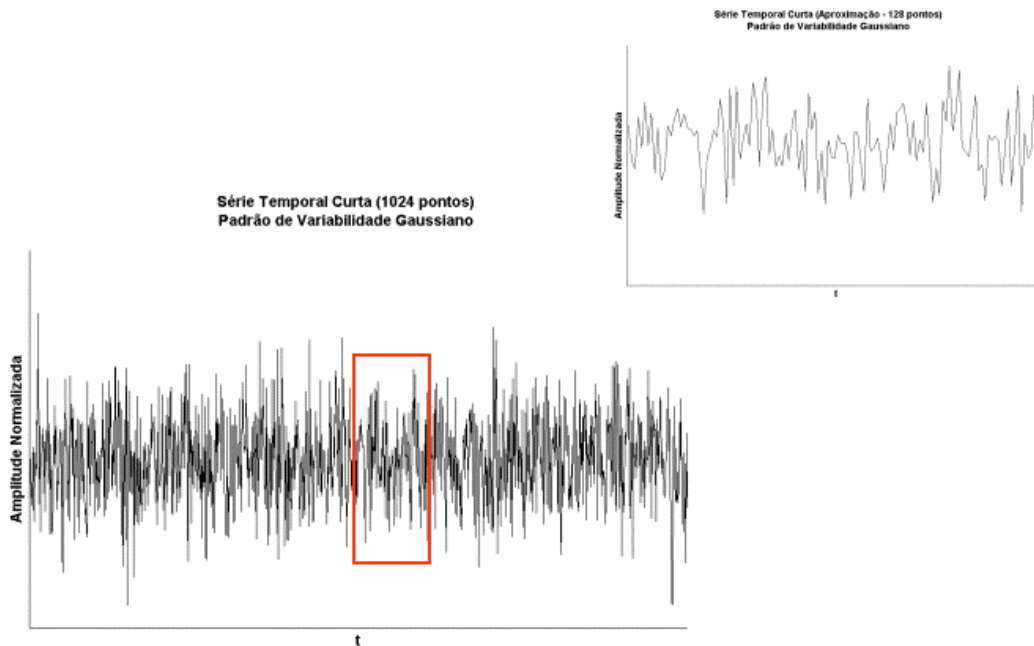


FIGURA 3.1 – Série temporal curta gaussiana (1024 pontos). No detalhe, uma sub-amostra com 128 pontos.

3.1.2 Padrão de Variabilidade Caótico 1D (Mapa Logístico)

Todo processo caótico compreende uma dinâmica determinística, porém com alto grau de imprevisibilidade devido a sua alta sensibilidade às condições iniciais e ao parâmetro de controle.

O primeiro padrão de variabilidade irregular do tipo caótico será gerado pelo Mapa Logístico, amplamente investigado na teoria dos sistemas dinâmicos caóticos (Peitgen et al., 1992). Originalmente, o mapa logístico foi introduzido para descrição de crescimento populacional (Verhulst, 1845), de acordo com a seguinte equação:

$$X_{t+1} = \lambda X_t (1 - X_t) \quad (3.1)$$

, onde X corresponde à fração populacional e λ (parâmetro de controle caótico) a uma constante positiva que depende das condições ecológicas.

No modelo original, assume-se que um limite máximo de população pode ser suportado pelo meio-ambiente. Caso a população ultrapasse este limite, pode ocorrer um desastre como, por exemplo, a rápida diminuição de alimentos, resultando na eliminação das espécies (Devaney, 1992).

Nas abordagens mais recentes o mapa logístico é utilizado para simular mapas acoplados que descrevem dinâmica de fluidos extensos e sistema de osciladores (Kaneko, 1990).

Na Figura 3.2 é apresentada uma ST constituída por 128 pontos que representa o padrão de variabilidade caótico 1D, gerado pelo mapa logístico, considerando $\lambda=4$. No detalhe, apresenta-se a ST curta constituída por 1024 pontos.

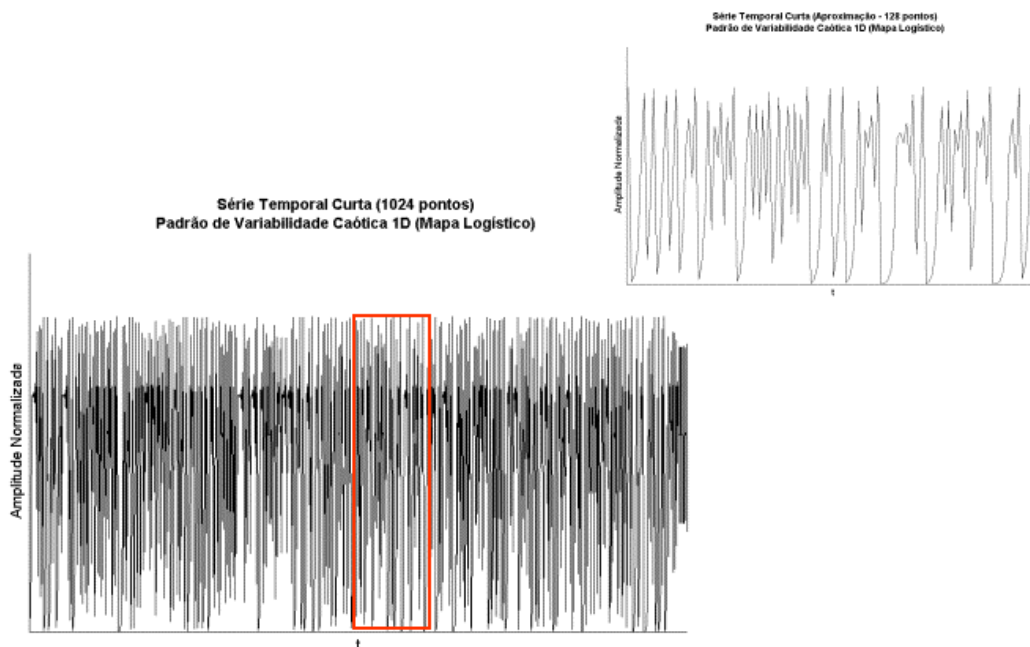


FIGURA 3.2 – Série temporal curta (1024 pontos) com padrão de variabilidade caótico 1D (Mapa logístico). No detalhe, uma sub-amostra com 128 pontos.

3.1.3 Padrão de Variabilidade Caótico 2D (Mapa de Hénon)

O Mapa de Hénon corresponde ao segundo padrão de variabilidade irregular do tipo caótico.

O mapa de Hénon foi introduzido por M. Hénon (Hénon, 1976) como simplificação de uma seção de Poincaré de equações diferenciais parciais. Trata-se de um dos exemplos mais estudados, devido ao comportamento caótico. Este mapa (quadrático e bidimensional) é representado na área de sistemas dinâmico pelas seguintes equações:

$$\begin{aligned} X_{t+1} &= Y_t + 1 - aX_t^2 \\ Y_{t+1} &= bX_t \end{aligned} \tag{3.2}$$

, onde as variáveis X e Y , representam pontos do plano cartesiano e os parâmetros canônicos a e b , definidos como 1.4 e 0.3 , respectivamente, apresentam um mapa com comportamento caótico. Para diferentes valores de a e b , o mapa pode apresentar comportamentos caótico, intermitente ou convergir para uma órbita periódica.

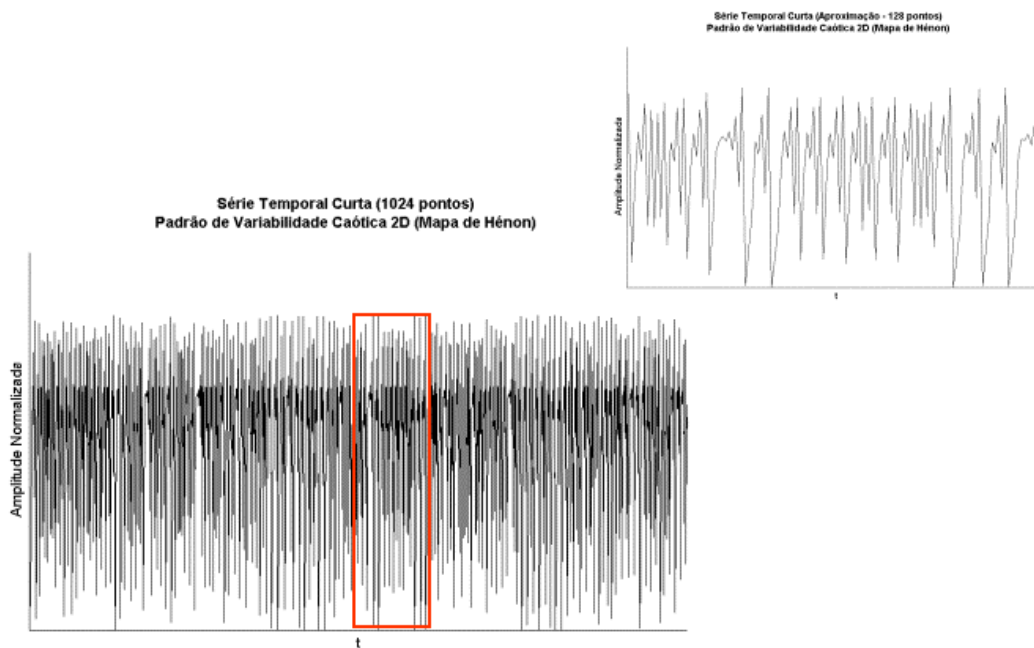


FIGURA 3.3 – Série temporal curta (1024 pontos) com padrão de variabilidade caótico 2D (Mapa de Hénon). No detalhe, uma sub-amostra com 128 pontos.

Na Figura 3.3 é apresentada uma ST constituída por 128 pontos que representa o padrão de variabilidade caótico 2D, gerado pelo mapa de Hénon com $a=1.4$ e $b=0.3$. No detalhe, apresenta-se a ST curta constituída por 1024 pontos.

3.1.4 Padrão de Variabilidade Caótico 3D (Modelo de Lorenz)

O terceiro modelo escolhido para representação de um padrão de variabilidade irregular é o Modelo de Lorenz. Na década de sessenta, o meteorologista americano Edward Lorenz estudava a evolução temporal da convecção atmosférica através de um modelo matemático tridimensional. Este modelo, conhecido como modelo de Lorenz, é representado pelo seguinte sistema de equações diferenciais parciais:

$$\begin{aligned}\frac{dX}{dt} &= \sigma(Y - X) \\ \frac{dY}{dt} &= X(\beta - Z) - Y \\ \frac{dZ}{dt} &= XY - \gamma Z\end{aligned}\tag{3.3}$$

, onde a variável X é proporcional à intensidade dos movimentos convectivos. Se $X > 0$, os movimentos ocorrem em sentido horário, caso contrário, se $X < 0$, os movimentos ocorrem em sentido anti-horário. A variável Y é proporcional à diferença de temperatura entre as correntes ascendentes e descendentes e Z é proporcional à distorção no perfil de temperatura vertical. Os parâmetros σ e β representam os números de Prandtl e Rayleigh, respectivamente, e γ é um fator geométrico (Lorenz, 1963).

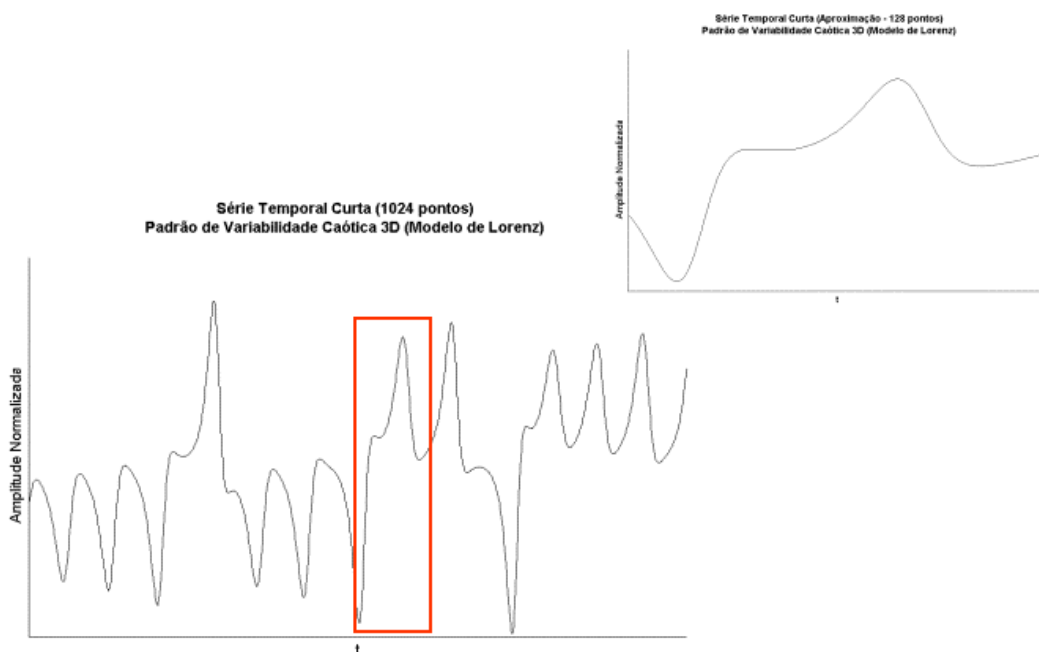


FIGURA 3.4 – Série temporal curta (1024 pontos) com padrão de variabilidade caótico 3D (Modelo de Lorenz). No detalhe, uma sub-amostra com 128 pontos.

A ST curta (1024 pontos), que representa o padrão de variabilidade caótico 3D gerado pelo modelo de Lorenz, é apresentada na Figura 3.4. Os parâmetros iniciais foram configurados da seguinte maneira: $\sigma=10$, $\beta=28$ e $\gamma=8/3$.

3.1.5 Padrão de Variabilidade Reativo-Difusivo (Equação de Kuramoto-Sivashinsky)

O próximo padrão de variabilidade irregular representado será do tipo reativo-difusivo, que pode ou não, conter uma componente caótica. Um processo reativo-difusivo apresenta resposta a instabilidades intrínsecas geradas pelo modelo. A classe de equações do tipo Ginzburg-Landau (Ginzburg e Landau, 1950) apresenta comportamento reativo-difusivo. Casos particulares são gerados pelas chamadas equações de amplitude, dentre as quais destaca-se a equação de Kuramoto-Sivashinsky que também possui comportamento caótico, porém modulado por uma das componentes espaciais.

A equação diferencial parcial de Kuramoto-Sivashinsky, largamente estudada na área de sistema dinâmicos para a modelagem de sistemas com comportamentos reativo-difusivo, pode ser representada da seguinte maneira (Rempel et al., 2004):

$$\partial_t u = -\partial_x^2 u - v \partial_x^4 u - \partial_x u^2 \quad (3.4)$$

onde v representa o parâmetro de viscosidade, e u é sujeito às condições limites periódicas $u(x, t) = u(x+2\pi, t)$. A equação foi nomeada após sua derivação como uma equação de fases da equação complexa de Ginzburg-Landau, apresentada por Kuramoto e Tsuzuky (Kuramoto e Tsuzuky, 1976) e um modelo para instabilidade hidrodinâmica, apresentado por Sivashinsky (Sivashinsky, 1977).

Na Figura 3.5 é apresentada uma ST que representa o padrão de variabilidade gerado pela equação de Kuramoto-Sivashinsky e a aproximação de uma sub-amostra da mesma constituída por 128 pontos.

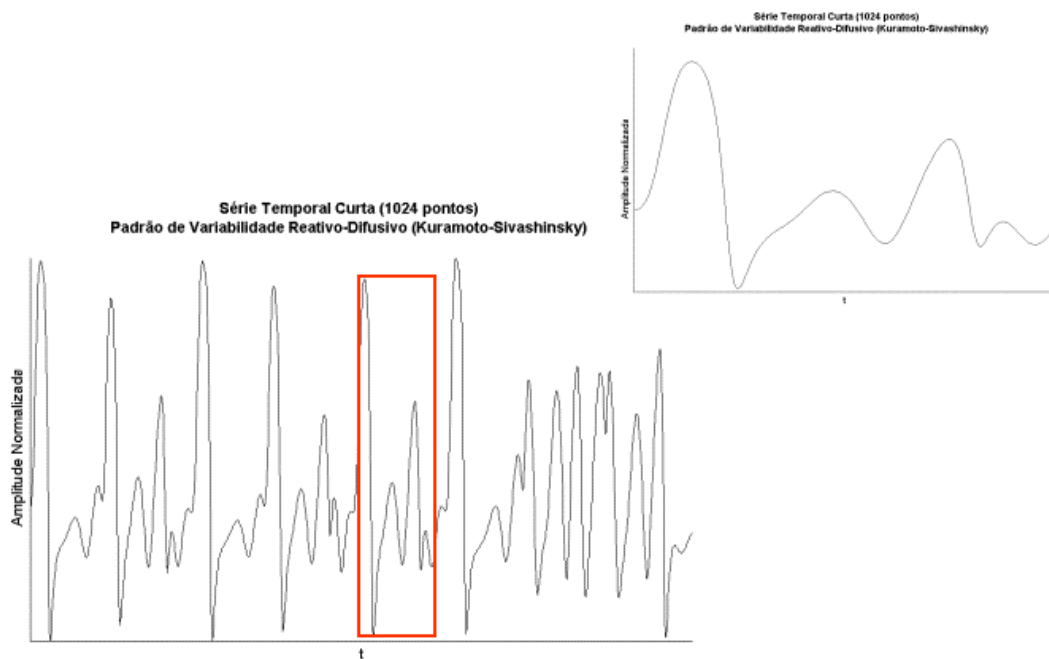


FIGURA 3.5 – Série temporal curta (1024 pontos) com padrão de variabilidade reativo-difusivo (Equação de Kuramoto-Sivashinsky). No detalhe, uma subamostra com 128 pontos.

Deve-se observar que para a escala pré-fixada de 1/1024 pontos, as frequências dos padrões gerados através de modelos discretos (Lorenz e Kuramoto-Sivashinsky) são mais altas que aquelas dos padrões gerados pelos sistemas contínuos (logístico e Hénon). Esse fato deve-se a influência das condições de contorno que tem papel preponderante na modulação dos padrões. Nos casos discretos não há restrições em relação aos contornos, o que reforça as oscilações locais elementares, não moduladas globalmente (Cross e Hohenberg, 1989; Rosa et al., 2000; Rosa et al., 2003).

3.1.6 Padrão de Variabilidade Turbulento (Dados de Velocidade do Vento – Projeto LBA)

O próximo padrão de variabilidade irregular corresponde à turbulência, caracterizado pela riqueza de escalas que resulta em um espectro de energias representado pelo

expoente de escala $-5/3$ (Kolmogorov, 1941). Os dados escolhidos para representar o padrão de variabilidade turbulento são de natureza experimental e foram obtidos através de medidas de velocidade do vento na floresta amazônica (Ramos et al., 2002).

Os dados de velocidade do vento, com comportamento turbulento, amostrados a 60 Hertz, foram obtidos através de uma campanha micro-meteorológica intensiva realizada pelo projeto LBA, durante os meses de janeiro a março de 1999. As medidas foram realizadas através de anemômetros sônicos tridimensionais (Campbell Scientific Inc., modelo CSAT-3-L60) (Figura 3.5) (Bolzan, 2002).

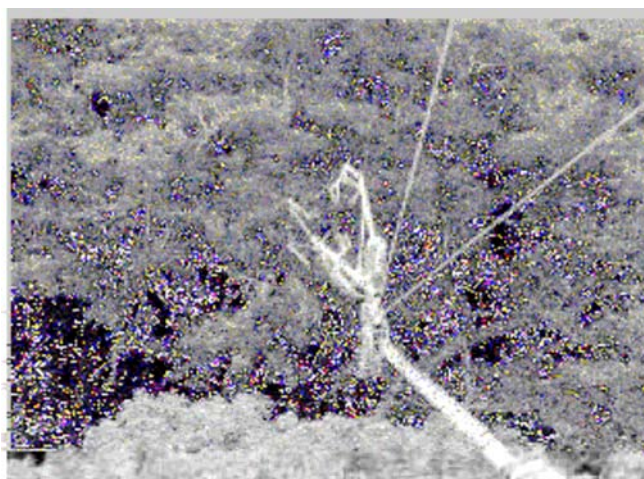


FIGURA 3.6 – Anemômetro Sônico tridimensional instalado na torre micrometeorológica da Reserva Biológica do Jaru.
FONTE: Bolzan (2002)

A ST curta (1024 pontos) apresentada na Figura 3.7 representa o padrão de variabilidade turbulento. No detalhe, é apresentada uma sub-amostra da mesma série constituída por 128 pontos.

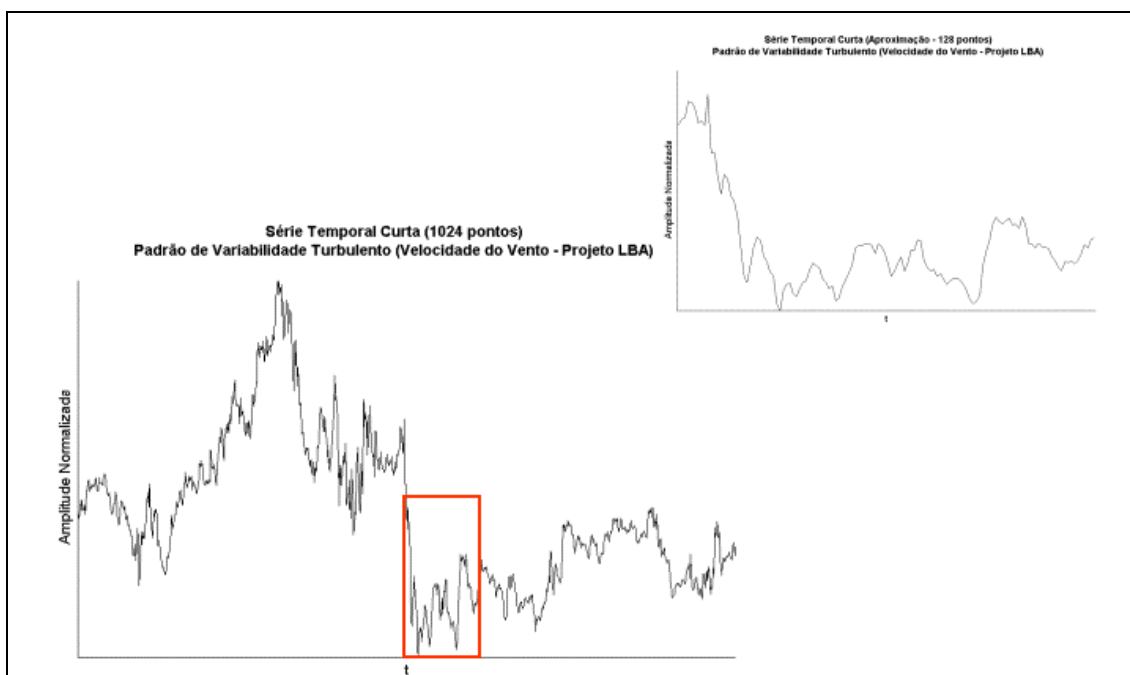


FIGURA 3.7 – Série temporal curta com padrão de variabilidade turbulento (Velocidade do Vento – projeto LBA) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.

3.2 Séries Temporais Ambientais

Nas próximas seções, serão apresentadas as ST que constituem os estudos de casos escolhidos. Todos os estudos de casos estão associados a medidas de variáveis atmosféricas, possuem caráter não-estacionário e são apresentados através de ST curtas.

3.2.1 Sistema Integrado de Monitoramento Ambiental

O Sistema Integrado de Monitoramento Ambiental (SIMA), projeto desenvolvido em conjunto entre a Universidade do Vale do Paraíba (UNIVAP) e o INPE, pode ser definido como um conjunto de hardware e software confeccionado para coleta e monitoramento de dados observados em sistemas hidrológicos (Stevenson et al., 1993).

A coleta de dados é realizada a partir de plataformas automáticas para coleta de dados (PCD) constituídas de um toróide (Figura 3.8), onde estão instalados: (i) sensores, (ii) sistema de armazenamento, (iii) bateria, (iv) painel solar e (v) antena de transmissão. Os dados são coletados em intervalos de tempo pré-programados desde janeiro de 2004, e transmitidos via satélite em tempo quase-real para um usuário que pode estar situado até 2500 Km distante do ponto de coleta (Stech et al., 2004).

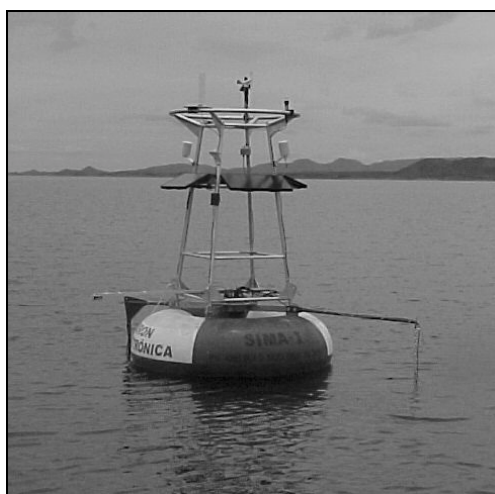


FIGURA 3.8 – Fotografia de uma PCD instalada no reservatório de Serra da Mesa.
FONTE: Stech et al. (2004)

O conjunto de variáveis monitoradas pelas PCD foi selecionado de acordo com os seguintes critérios (Stech et al., 2004):

- Relevância para a caracterização dos ambientes aquáticos;
- Relevância como indicador de impacto ambiental (variáveis que respondem de maneira consistente às alterações no funcionamento do sistema aquático);
- Relevância no processo de emissão de gases do efeito estufa;
- Viabilidade técnica de obtenção de medidas através das PCD.

Com base nesses critérios, o sistema de medição automática monitora as seguintes variáveis ambientais: clorofila, PH, turbidez, concentração de O₂, condutividade elétrica, nitrato, amônia, temperatura da água, pressão atmosférica, umidade relativa,

temperatura do ar, direção e intensidade do vento, radiação solar, fluxo e direção da água.

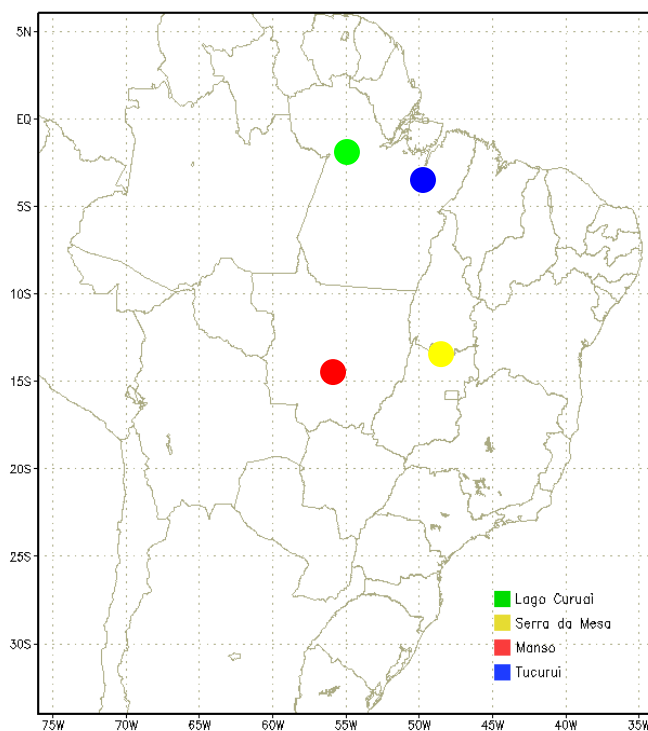


FIGURA 3.9 – Localização dos reservatórios selecionados para a realização do projeto SIMA.

A área de estudo do SIMA corresponde às regiões Norte e Centro-Oeste do Brasil (Figura 3.9). Três PCD foram instaladas nos reservatórios de Tucuruí e Serra da Mesa, localizados na bacia Araguaia-Tocantins (região de floresta tropical e savana) e Manso, localizado ao Norte da bacia do Paraná (região de transição entre savana e floresta tropical). A última PCD foi instalada no lago Curuai, localizado no baixo planalto Amazônico (região de floresta tropical) (Stech et al., 2004).

As Figuras 3.10 a 3.11 apresentam dois exemplos de ST de temperaturas do ar observadas nos reservatórios de Manso e Serra da Mesa, respectivamente.

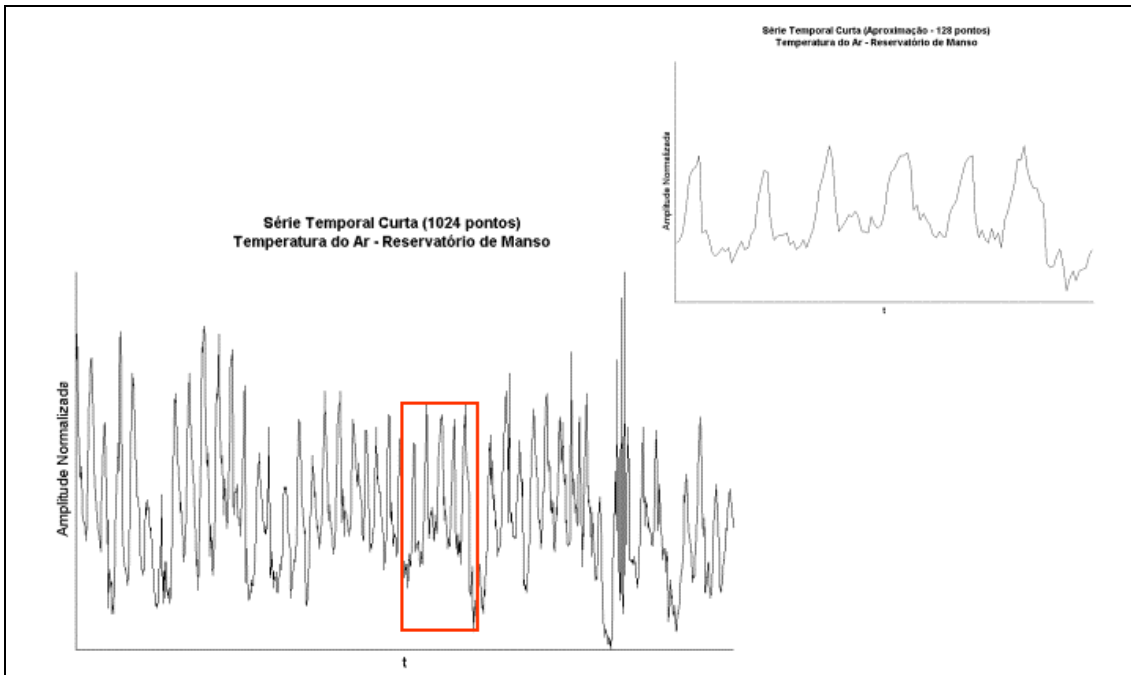


FIGURA 3.10 – Série temporal curta de temperatura do ar (reservatório de Manso) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.

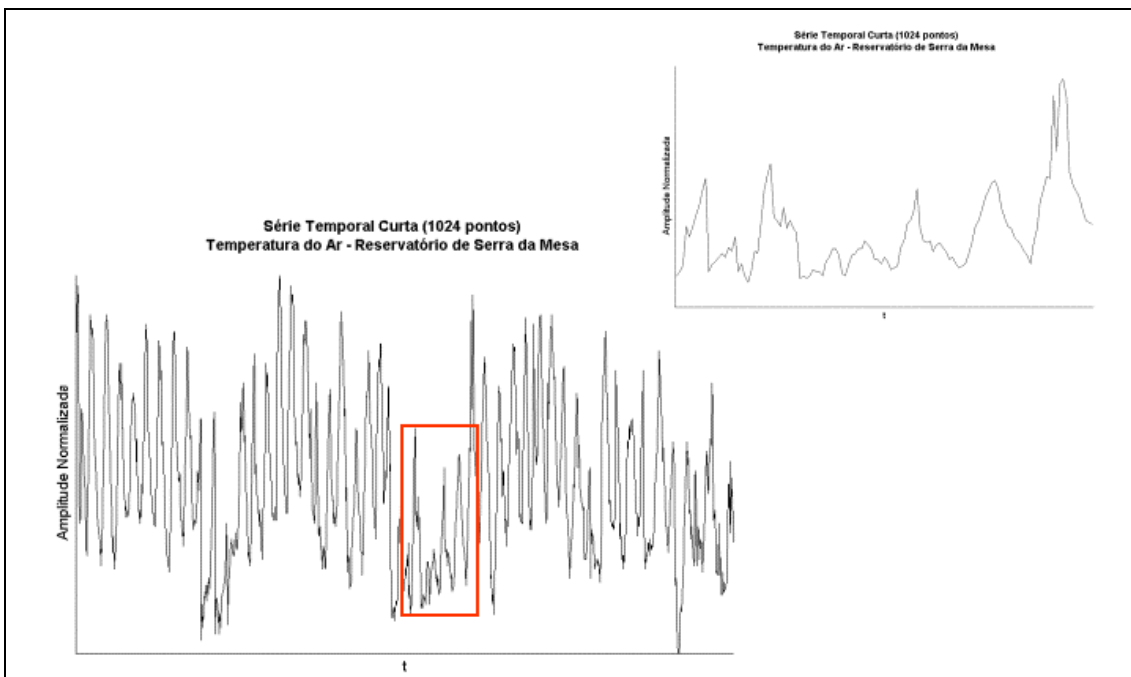


FIGURA 3.11 – Série temporal curta de temperatura do ar (reservatório de Serra da Mesa) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.

3.2.2 Dados Oceanográficos

As séries oceanográficas analisadas consistem um conjunto de medidas de amplitudes de correntes oceânicas (componentes Leste e Norte), ainda de caráter preliminar, gentilmente fornecidas pela empresa *MV – Oceânica*.

É interessante observar que, embora os processos subjacentes tenham as mesmas características, a série da componente leste (Figura 3.12) apresenta uma pequena modulação em baixa frequência mais acentuada se comparada com a componente norte (Figura 3.13). Trata-se de um exemplo de detalhe, por suposição através da inspeção visual dos padrões de variabilidade, que pode ser caracterizado a partir de técnicas apropriadas para análise de ST.

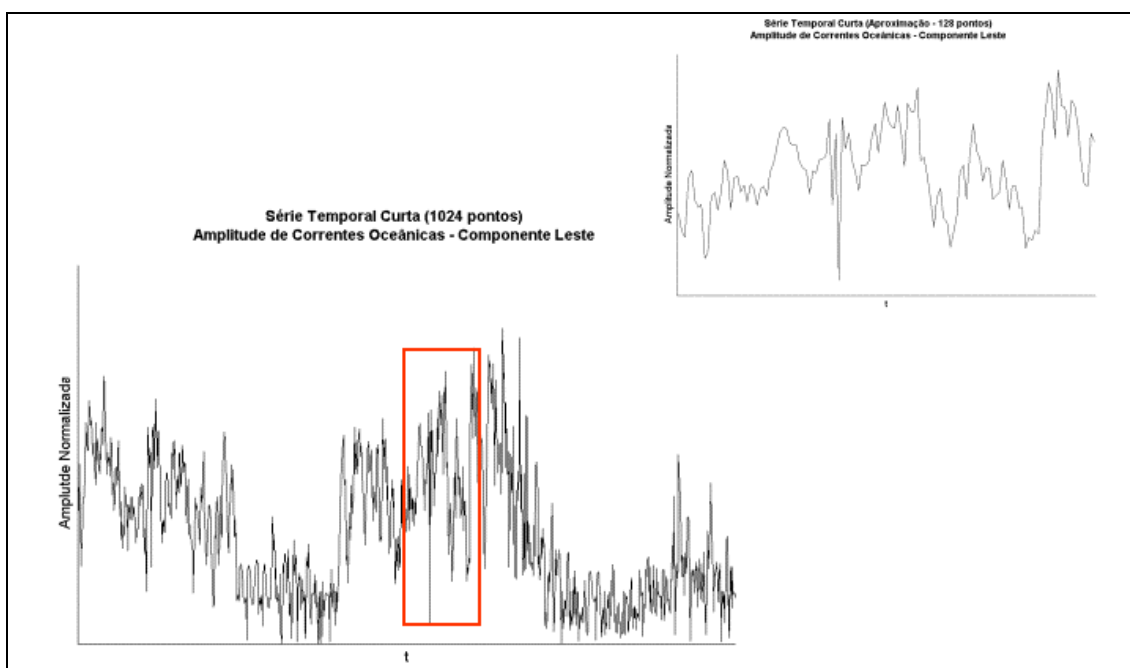


FIGURA 3.12 – Série temporal curta de amplitudes de correntes oceânicas (Componente Leste) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.

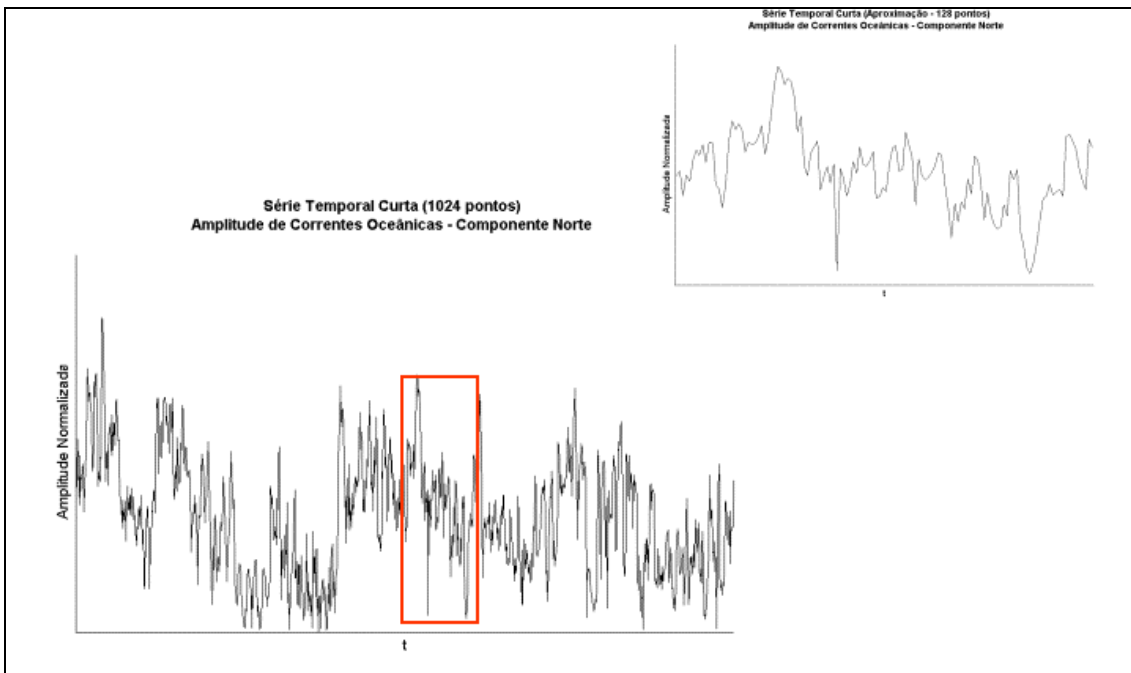


FIGURA 3.13 – Série temporal curta de amplitudes de correntes oceânicas (Componente Norte) (1024 pontos). No detalhe, uma sub-amostra de 128 pontos.

CAPÍTULO 4

TÉCNICAS ESTATÍSTICAS EM SCDP *SCILAB*

Nas próximas seções serão realizadas introduções sobre softwares de domínio público e sobre o ambiente de desenvolvimento *Scilab*. Em seguida (seção 4.3), serão apresentados alguns resultados obtidos através da aplicação das técnicas de transformada de ondeletas e de GPA1D, implementadas em ambiente *Scilab*, nos conjuntos de dados descritos no Capítulo 3. Serão realizadas análises ilustrativas das séries com o único objetivo de validar os algoritmos desenvolvidos em SCDP *Scilab*.

4.1 Softwares de Domínio Público

O conceito de software de domínio público, introduzido pela *Free Software Foundation* (FSF) no início da década de oitenta, tem apresentado uma crescente popularidade nos últimos anos. Sua idéia central consiste em garantir aos usuários liberdade para acessar e estudar o código fonte dos programas e, caso seja necessário, alterá-los e, conseqüentemente, disseminá-los livremente para a comunidade.

Inicialmente sem muita visibilidade, a comunidade de software de domínio público tem crescido extensivamente com a colaboração de voluntários interagindo a partir da Internet. Além disso, contribui para seu crescimento o poder de propaganda e a forma de comercialização gratuita, além da alta qualidade oferecida por produtos de softwares como, por exemplo, o Linux.

O sistema operacional Linux, desenvolvido no início da década de noventa na Finlândia, é o exemplo de maior sucesso no contexto de softwares de domínio público. Distribuído livremente através da Internet, seguindo as diretrizes da *GNU (GNU is Not Unix) Public License* (GPL) (www.gnu.org), o Linux é considerado um sistema operacional

maduro e tem sido utilizado e desenvolvido, tanto em ambiente acadêmico como em instituições públicas e privadas.

No entanto, a adoção de softwares de domínio público por instituições não envolve apenas as questões tecnológicas como segurança (códigos fechados podem ocultar vírus), autonomia (para realizar alterações nos programas) e praticidade (facilidade de uso e modificações). Em geral, a questão econômica também é avaliada e considerada relevante para a tomada de decisões, uma vez que os softwares de domínio público consomem menos investimentos financeiros.

A Internet é considerada um vasto e rico repositório de softwares de domínio público. Implementações de alta qualidade podem ser facilmente encontradas para diferentes finalidades tais como, matemática, análise numérica, computação paralela, processamento de dados, visualização de imagens, preparação de textos e apresentações, dentre outras.

No contexto de softwares de domínio público para a computação científica, diversas opções podem ser destacadas. Alguns exemplos de sucesso são: (i) *Scilab*, um ambiente de desenvolvimento de softwares destinado à resolução de problemas numéricos; (ii) Octave, uma linguagem de programação de alto-nível destinada à utilização de computações numéricas e (iii) R, uma linguagem de programação voltada para análise estatística computacional e gráfica.

Para a realização do presente trabalho optou-se pela utilização do ambiente de desenvolvimento *Scilab* devido a algumas características relevantes que destacam suas funcionalidades: (i) distribuição livre a partir da Internet com o código fonte aberto; (ii) desenvolvimento contínuo realizado por uma equipe de desenvolvimento; (iii) sistema robusto para auxílio ao usuário; (iv) linguagem de programação simples e de fácil aprendizado; (v) facilidade para definição de funções e (vi) interface com funções escritas em linguagens de programação Fortran e C.

4.2 Software Científico de Domínio Público: *Scilab*

Criado em 1990, por um grupo de pesquisadores do *Institut de Recherche en Informatique et en Automatique* (INRIA) e da *École Nationale des Ponts et Chaussées* (ENPC), na França, o *Scilab* é definido como um ambiente de desenvolvimento de softwares destinado à resolução de problemas numéricos (Pires, 2002).

Desde 1994 o *Scilab* é distribuído livremente a partir da Internet, seguindo o conceito de software de domínio público. Existem diferentes distribuições desse ambiente de desenvolvimento, e através de uma rápida busca no sítio do *Scilab* na Internet (www.scilab.org), pode-se obter o código fonte do programa, bem como, versões pré-compiladas para diversos sistemas operacionais (Anexo C).

A versão 3.1.1 (Figura 4.1), pré-compilada para sistema operacional Windows 2000, foi utilizada durante o desenvolvimento deste trabalho.

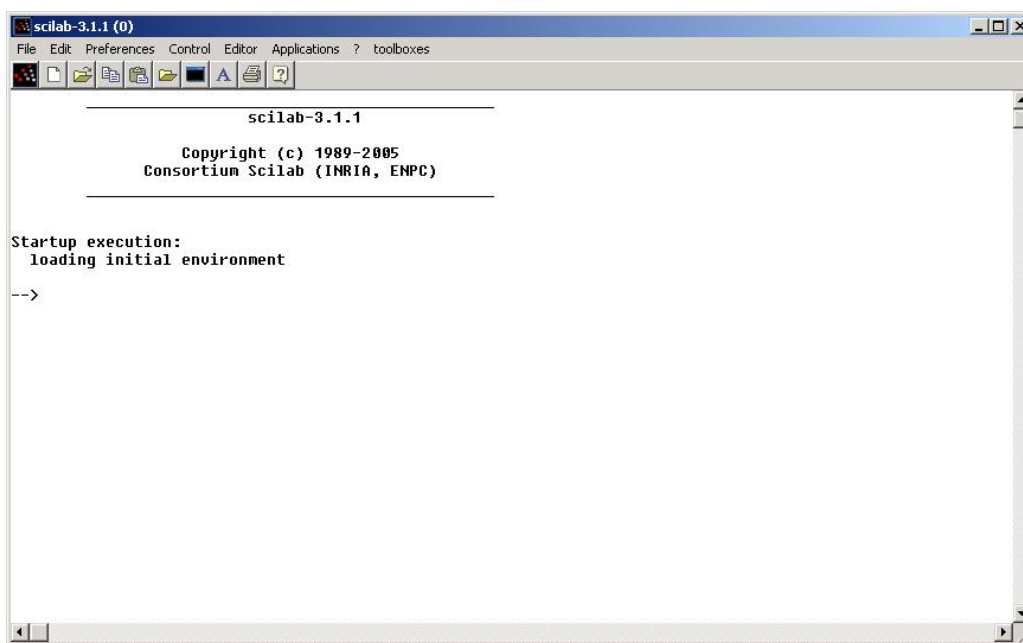


FIGURA 4.1 – Janela inicial do *Scilab* 3.1.1, versão pré-compilada para Windows 2000.

A partir de 2003, o *Scilab* passou a ser mantido por um consórcio de empresas e instituições francesas denominado “Consórcio *Scilab*”, cujos principais objetivos são:

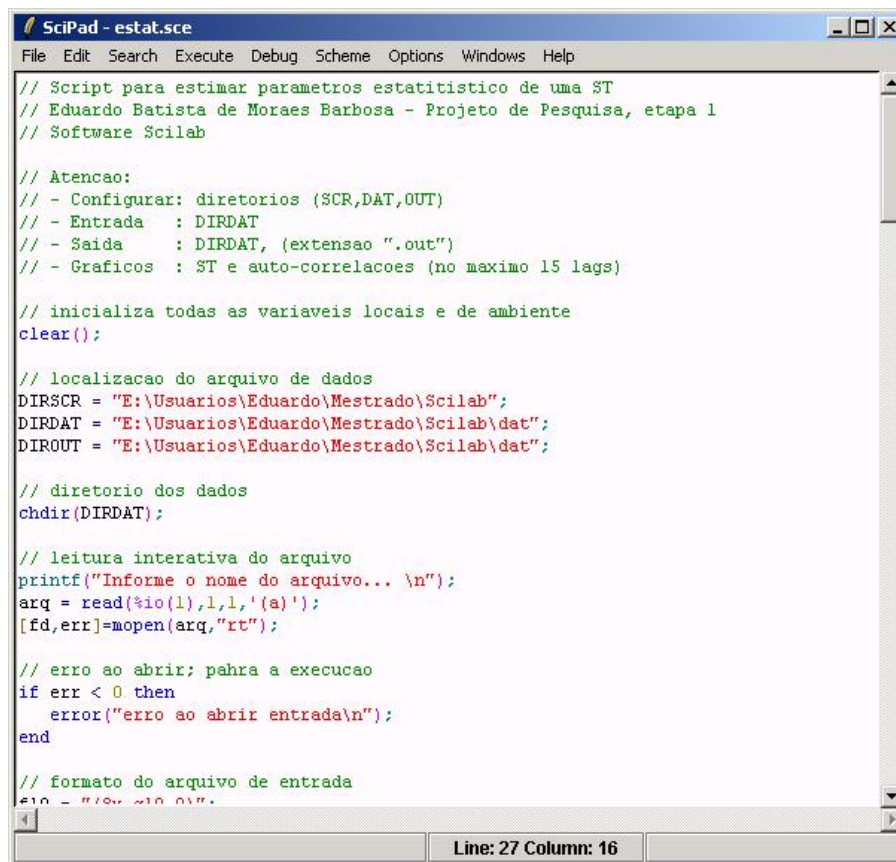
- Organização e intercâmbio entre a equipe de desenvolvimento, visando incorporar ao *Scilab* os últimos avanços científicos na área de computação numérica;
- Organização e intercâmbio entre usuários, com a finalidade de tornar a utilização do ambiente de desenvolvimento mais efetiva na indústria, na educação e na pesquisa;
- Angariar recursos para a manutenção da equipe de desenvolvimento, visando garantir suporte adequado às necessidades dos usuários.

Adicionalmente, o *Scilab* possui algumas características semelhantes ao MATLAB® (www.mathworks.com), isto é, (i) utilização de ambiente gráfico combinado com linguagem de programação de alto nível; (ii) implementação de funções para manipulação de matrizes e (iii) suporte ao desenvolvimento de conjuntos de funções destinadas a aplicações específicas (Pires, 2002).

4.2.1 Descrição do Funcionamento do *Scilab*

A programação *Scilab* oferece grande parte das estruturas de controle de fluxo e *loop* encontradas em linguagens convencionais. No entanto, as principais diferenças são que o *Scilab* é um interpretador de comandos e por isso não há necessidade de declarações prévias dos tipos das variáveis que serão utilizadas ao longo do programa.

Os programas, também chamados de *scripts*, são implementados em arquivos textos contendo seqüências de comandos *Scilab*, sem acentuação. Toda a parte de implementação, bem como, execuções dos programas podem ser realizadas com o auxílio das opções oferecidas pelo *Scipad*, um editor de códigos com sintaxe destacada e interface amigável (Figura 4.2) incorporado ao *Scilab*.



```
SciPad - estat.sce
File Edit Search Execute Debug Scheme Options Windows Help

// Script para estimar parametros estatistico de uma ST
// Eduardo Batista de Moraes Barbosa - Projeto de Pesquisa, etapa 1
// Software Scilab

// Atencao:
// - Configurar: diretorios (SCR,DAT,OUT)
// - Entrada   : DIRDAT
// - Saida    : DIRDAT, (extensao ".out")
// - Graficos : ST e auto-correlacoes (no maximo 15 lags)

// inicializa todas as variaveis locais e de ambiente
clear();

// localizacao do arquivo de dados
DIRSCR = "E:\Usuarios\Eduardo\Mestrado\Scilab";
DIRDAT = "E:\Usuarios\Eduardo\Mestrado\Scilab\dat";
DIROUT = "E:\Usuarios\Eduardo\Mestrado\Scilab\dat";

// diretorio dos dados
chdir(DIRDAT);

// leitura interativa do arquivo
printf("Informe o nome do arquivo... \n");
arq = read(%io(1),1,1,'a');
[fd,err]=mopen(arq,"rt");

// erro ao abrir; pabra a execucao
if err < 0 then
    error("erro ao abrir entrada\n");
end

// formato do arquivo de entrada
%io = "%s %s %s";

Line: 27 Column: 16
```

FIGURA 4.2 – *Scipad* editor de programas fornecido em conjunto com o *Scilab*.

Por convenção, programas *Scilab* possuem a extensão *.sce* e são executados via *prompt* de comandos, com o auxílio do comando *exec*. Por outro lado, diferente dos programas, as funções *Scilab* possuem a extensão *.sci* e estruturas pré-determinadas, cujas principais características são:

- As variáveis definidas no escopo da função, chamadas de variáveis locais, não permanecem no ambiente *Scilab* após a execução da função;
- As entradas e saídas das funções são claramente definidas;
- Uma função, após ser definida, pode ser chamada a qualquer momento.

Além disso, as funções devem respeitar as seguintes definições: iniciar com a palavra-chave *function*, definir claramente as variáveis de saída, o nome da função e as variáveis

de entrada, e terminar com a palavra-chave *endfunction*. Toda função deve ser definida pelo comando *getf* e pode ser executada a partir de chamadas de seu nome por outras funções ou via linha de comando.

4.3 Resultados Preliminares

Nas próximas seções serão apresentados alguns resultados preliminares obtidos com a aplicação de técnicas estatísticas, implementadas em *Scilab*, (i) convencionais e (ii) avançadas (transformada de ondeletas e GPA1D) nos conjuntos de dados descritos no Capítulo 3.

4.3.1 Estatística Convencional

Os parâmetros estatísticos convencionais (média, variância e desvio-padrão) foram obtidos a partir de funções intrínsecas (*mean*, *variance* e *st_deviation*, respectivamente) oferecidas pelo SCDP *Scilab*. Em geral, os parâmetros estatísticos convencionais são utilizados como uma apresentação resumida dos dados de ST.

4.3.1.1 Média, Variância e Desvio-Padrão

TABELA 4.1 – Parâmetros estatísticos calculados para dois conjuntos de dados atmosféricos em *Scilab*.

Conjunto de Dados	\bar{x}	σ^2	σ
Temperatura do ar (Serra da Mesa)	27.51	8.88	2.98
Amplitudes de correntes oceânicas (componente Leste)	8.70	124.25	11.15

Na Tabela 4.1 são apresentados os resultados dos parâmetros estatísticos convencionais (média, variância e desvio-padrão) calculados para dois diferentes conjuntos de dados

atmosféricos (temperatura do ar no reservatório de Serra da Mesa e amplitudes de correntes oceânicas).

4.3.1.2 Autocorrelação

O comportamento de duas variáveis que compõem um conjunto de dados pode ser observado graficamente a partir de um diagrama denominado diagrama de dispersão (Figura 4.3) e medido através do coeficiente de correlação (Capítulo 2), ou grau de dependência entre as variáveis. Quando o objetivo consiste em medir a dependência de uma variável com ela mesma em diferentes passos de tempo, o processo é, então, chamado de autocorrelação.

Na Figura 4.3 são apresentados os diagramas de dispersão para diferentes conjuntos de dados de ciências atmosféricas. Linhas de tendências foram ajustadas aos dados dos diagramas, com a finalidade de auxiliar a classificação do tipo de autocorrelação obtida. As autocorrelações foram calculadas considerando-se *lag* 1 e, em todos os casos, observou-se linhas com inclinação da esquerda para direita, que indicam autocorrelação positiva. Esta informação pode ser confirmada ao se observar os resultados da Tabela 4.2.

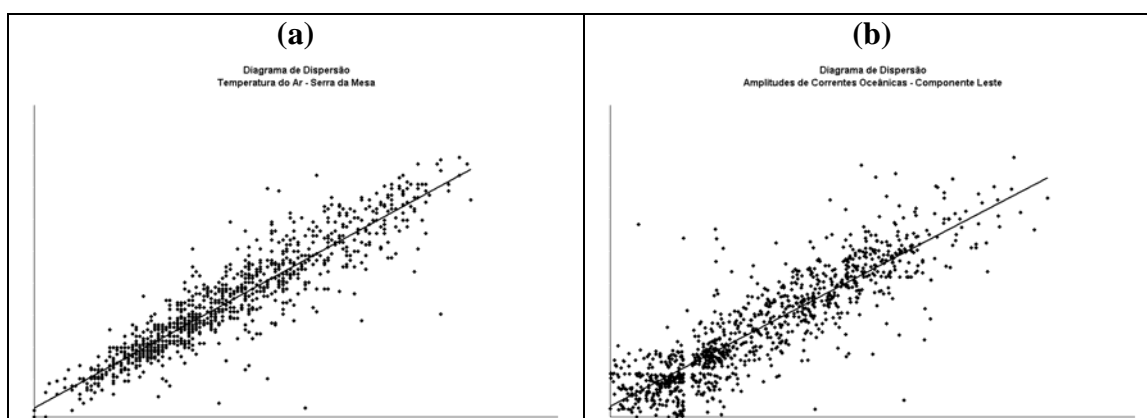


FIGURA 4.3 – Diagramas de dispersão, em *Scilab*, para dois conjuntos de dados atmosféricos. (a) Temperatura do ar (Serra da Mesa); (b) Amplitudes de correntes oceânicas (componente Leste).

TABELA 4.2 – Autocorrelações de diferentes conjuntos de dados atmosféricos.

Conjuntos de Dados	ρ (%)
Temperatura do ar (Serra da Mesa)	0.92
Amplitudes de correntes oceânicas (componente Leste)	0.87

4.3.2 Implementação de Técnicas Estatísticas Avançadas em Scilab

Nas próximas seções serão apresentados e comparados os resultados obtidos através da aplicação das técnicas de transformada de ondeletas e de GPA1D, implementadas em SCDP *Scilab* e softwares proprietários (MATLAB® e IDL), em diferentes conjuntos de dados apresentados no Capítulo 3.

4.3.2.1 Transformada de Ondeletas em Scilab

A transformada de ondeletas foi adaptada ao ambiente de programação *Scilab* com o auxílio de um conjunto de programas fornecidos livremente pela biblioteca de softwares Wavelab (www-stat.stanford.edu/~wavelab/), originalmente, desenvolvida para o ambiente de programação MATLAB®. A Wavelab pode ser definida como uma biblioteca de softwares destinada à análise de ondeletas. Inicialmente utilizada por universidades Americanas, esta biblioteca possui um conjunto de programas que viabilizam o aprendizado de alguns aspectos práticos da análise de ondeletas como, por exemplo, a reprodução de figuras já publicadas por alguns autores.

Como salientado anteriormente, o *Scilab* possui algumas características semelhantes ao MATLAB®. Entretanto, durante o desenvolvimento deste trabalho, utilizando a versão 3.1.1 do *Scilab*, observou-se à necessidade de implementações e adaptações de algumas funções da Wavelab (originalmente desenvolvidas em MATLAB®), para o ambiente *Scilab*.

Após a fase de implementações foram realizados diversos testes, com a finalidade de comparar os resultados da transformada de ondeletas implementada em ambiente *Scilab* com a implementação original, em MATLAB®.

Nas Figuras 4.4 e 4.5, o gráfico superior (versão *Scilab*) representa a ST original (referente aos dados analisados). O gráfico inferior à esquerda (versão *Scilab*) representa o periodograma obtido a partir da transformada de ondeletas, onde o eixo x é a escala temporal, em dias, e o eixo y representa os períodos da série. A linha contínua no periodograma, obtida na versão MATLAB®, é chamada cone de influência e os períodos fora deste cone devem ser negligenciados por não possuírem confiança estatística. A barra de cores representa a energia associada a cada período da série.

Adicionalmente, adaptações nos programas originais possibilitaram a obtenção do espectro global de ondeletas (*Global Wavelet Spectrum*) através do cálculo das variâncias (energias) associadas a cada um dos períodos da série. Este procedimento consiste na somatória das energias de toda energia associada a cada período. O gráfico à direita (versão *Scilab*) representa o espectro de ondeleta global (GWS), geralmente, utilizado na identificação de períodos associados à maior energia (Torrence e Compo, 1998; Bolzan, 2004).

Na Figura 4.4 são apresentados os resultados da transformada de ondeletas aplicada em um conjunto de dados de temperatura do ar no reservatório de Serra da Mesa, amostrados em frequência horária, desenvolvidos em *Scilab* e MATLAB®, para validação do primeiro.

Através de observações nos gráficos da Figura 4.4 (a e b), pode-se verificar que os resultados obtidos são semelhantes para as duas implementações comparadas, isto é, *Scilab* (Figura 4.4-a) e MATLAB® (Figura 4.4-b). Entretanto, destaca-se que no periodograma confeccionado pelo *Scilab* (Figura 4.4-a) existem limitações relacionadas à resolução das cores. No entanto, esta limitação não impede a identificação dos

períodos relevantes da série. Destaca-se ainda neste periodograma, a ausência da linha contínua que representa o cone de influência, observada após a geração das imagens para este documento.

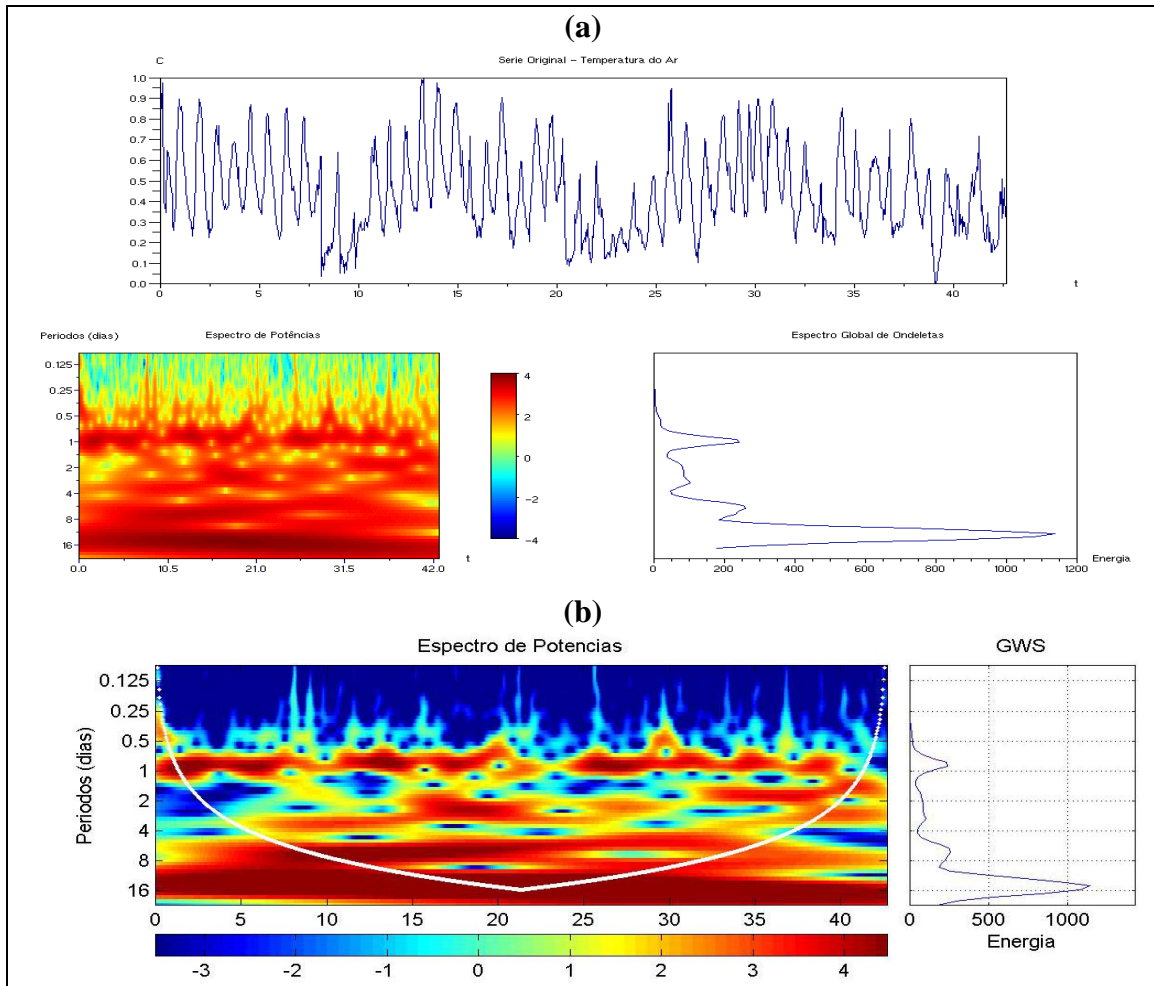


FIGURA 4.4 – Resultados da transformada de ondeletas aplicada em dados de temperatura do ar no reservatório de Serra da Mesa. (a) *Scilab*; (b) *MATLAB*®

Através das análises no periodograma da Figura 4.4 pode-se identificar, durante toda a ST, um sinal persistente em escala de aproximadamente 12 horas, possivelmente, associado ao período diurno de dados. No gráfico de GWS é possível identificar este ciclo, bem como, um período de energia intensa em escala de 8 dias, que pode estar relacionado com a passagem de sistemas convectivos sobre o reservatório.

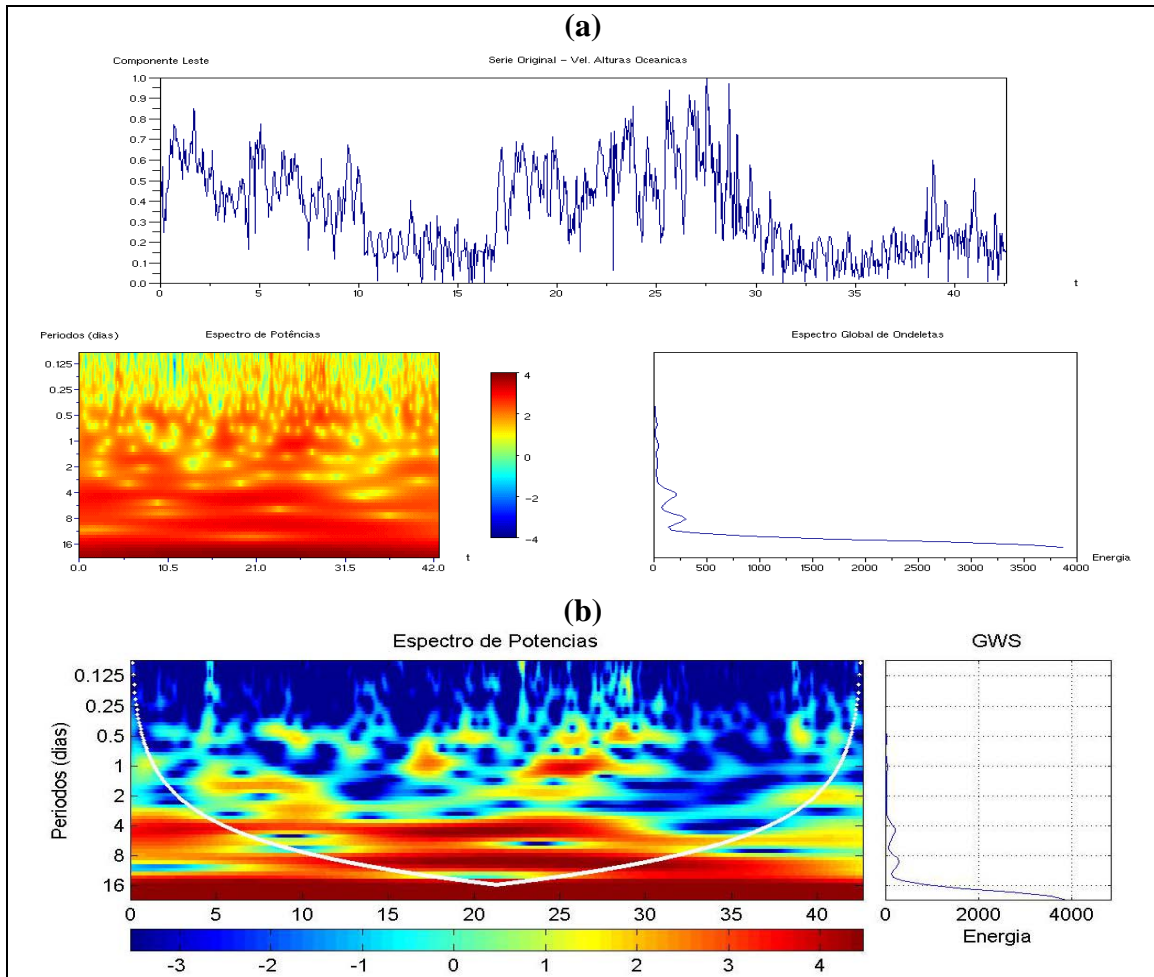


FIGURA 4.5 – Resultados da transformada de ondeletas aplicada em dados de amplitudes de correntes oceânicas (componente leste). (a) *Scilab*; (b) *MATLAB*®

Com o auxílio do periodograma e do gráfico de GWS (Figura 4.5) pode-se observar que as amplitudes das correntes oceânicas (componente leste) apresentam-se mais intensa em escala de aproximadamente 12 horas, possivelmente, associado ao ciclo das marés.

4.3.2.2 Análise de Padrões-Gradiente em *Scilab*

No contexto da análise de ST curtas, dada à falta de robustez estatística da amostra, até o final da década de 90 não havia sido definido um paradigma para sustentar a busca de uma metodologia para caracterização desse tipo de série. Em 2002, Assireu et al. (2002), baseado no trabalho de Rosa et al. (1999), introduziram uma nova proposta, denominada GPA1D, para análises de ST não-estacionárias curtas:

“Para um dado processo dinâmico subjacente, todo o conjunto de possíveis padrões de variabilidade no tempo resultante do mesmo, deve ser caracterizado por uma faixa restrita e única de valores, calculados a partir da relação entre a escala global e escala local de variabilidade da série”.

Isto é, tomando como exemplo um processo caótico, caso seja possível calcular essa relação, seus valores definirão uma faixa típica para o padrão de variabilidade gerado por processos caóticos de tal forma que, para uma dada série cujo processo subjacente é desconhecido, será possível inferir qual processo (nesse caso, caótico ou não) é candidato a explicar o padrão observado. Esta abordagem é de suma importância para variáveis cujas séries de medidas, usualmente, estão restritas a uma seqüência de N pontos que não satisfazem a validade estatística ($N \gg 10^3$).

A técnica GPA1D foi adaptada ao ambiente de programação *Scilab* com o auxílio de um extenso conjunto de programas desenvolvido pelos pesquisadores do Laboratório Associado de Computação e Matemática Aplicada (LAC/INPE), originalmente, para o ambiente de programação IDL.

Diferentemente da transformada de ondeletas, não foram identificadas semelhanças entre as sintaxes de programação IDL e *Scilab*, por isso a necessidade de adaptar todos os programas originais para o ambiente *Scilab*.

Após a fase de implementações foram realizados diversos testes, com a finalidade de comparar os resultados da análise de padrões-gradiente implementada em *Scilab* com a implementação original, em IDL.

Nas Tabelas 4.3 e 4.4 são apresentados os valores médios do primeiro momento gradiente ($\langle g_1^a \rangle$) obtidos com a aplicação da GPA1D em dois conjuntos de dados canônicos (caótico 3D e turbulento).

TABELA 4.3 – Resultados da GPA1D para o padrão de variabilidade caótico 3D (modelo de Lorenz).

	<i>Scilab</i>	IDL
$\langle g_1^a \rangle$	1.000	1.000

Na Tabela 4.3 são apresentados os resultados obtidos com a aplicação da GPA1D no conjunto de dados com padrão de variabilidade caótico 3D (modelo de Lorenz). Pode-se observar a partir desta que os resultados são idênticos para as duas implementações em análise, isto é, *Scilab* e IDL. A partir destes resultados observou-se que os valores de $\langle g_1^a \rangle$ pertencem ao intervalo $\langle g_1^a \rangle \pm \sigma = 1.001 \pm 0.002$, sugerido por Assireu et al. (2002) para a caracterização de padrões quase-periódicos.

Na Tabela 4.4 são apresentados os resultados obtidos para o conjunto de dados com padrão de variabilidade turbulento (velocidade do vento – projeto LBA). Como observado anteriormente (Tabela 4.3), estes resultados (Tabela 4.4) também são idênticos para as duas implementações (*Scilab* e IDL).

Os resultados apresentados (Tabela 4.4) indicam que os valores de $\langle g_1^a \rangle$ pertencem ao intervalo $\langle g_1^a \rangle \pm \sigma = 1.111 \pm 0.002$, sugerido por Assireu et al. (2002) para a caracterização de padrões turbulentos.

TABELA 4.4 – Resultados da GPA1D para o padrão de variabilidade turbulento (velocidade do vento – projeto LBA).

	<i>Scilab</i>	IDL
$\langle g_1^a \rangle$	1.111	1.111

No próximo capítulo será apresentada uma metodologia alternativa, que utiliza conjuntamente as decomposições por ondeletas e GPA1D, em SCDP *Scilab*, para a caracterização do padrão de variabilidade de séries curtas e não-estacionárias, que denominaremos por *Espectro de Assimetrias*.

CAPÍTULO 5

APLICAÇÃO CONJUNTA DA TRANSFORMADA DE ONDELETAS E ANÁLISE DE PADRÕES-GRADIENTE: ESPECTRO DE ASSIMETRIAS

Nas próximas seções serão apresentadas as técnicas de decomposição por ondeletas e de GPA1D, que juntas irão compor uma metodologia alternativa para análise de ST curtas e não-estacionárias. Esta metodologia, denominada *espectro de assimetrias*, consiste na aplicação conjunta da decomposição por ondeletas e GPA1D em uma ST qualquer, com a finalidade de caracterizar o seu padrão de variabilidade.

5.1 Decomposição Tempo-Freqüência por Transformada de Ondeletas

As funções ondeletas podem ser do tipo discreto (Haar, 1910) ou do tipo contínuo (Morlet, 1983). As ondeletas discretas, em geral, são utilizadas na decomposição e filtragem de ST. Neste trabalho cada decomposição será considerada uma ST, cujo padrão de variabilidade será caracterizado por sua assimetria correspondente. Após testes com transformadas contínuas e discretas, optamos pela segunda, uma vez que esta não provoca redundâncias nas freqüências geradas e o tratamento do sinal temporal considera o comprimento da ordem de potência de dois mais próxima, também referenciado como *diádico*. Ou seja, $N = 2^d$, onde N é o comprimento da ST e d o número de freqüências possíveis para as decomposições.

No contexto das séries curtas, as amostras consideradas serão constituídas por 1024 pontos, o que fornece uma decomposição com 10 freqüências, seguindo o critério diádico. Testes com as ondeletas discretas de Haar e de Daubechies (Daubechies, 1988) evidenciaram que a segunda preserva a suavidade do padrão de variabilidade em cada freqüência, enquanto que, a primeira, devido a sua descontinuidade inerente, impõe singularidades indesejáveis, ocasionando perda de informação em relação aos padrões

de variabilidade. As Figuras 5.1 e 5.2 apresentam duas decomposições da série gaussiana, sendo uma realizada com a transformada de Haar (Figura 5.1) e outra com a transformada de Daubechies (Figura 5.2). Em ambas decomposições, o número original de pontos é preservado em cada frequência.

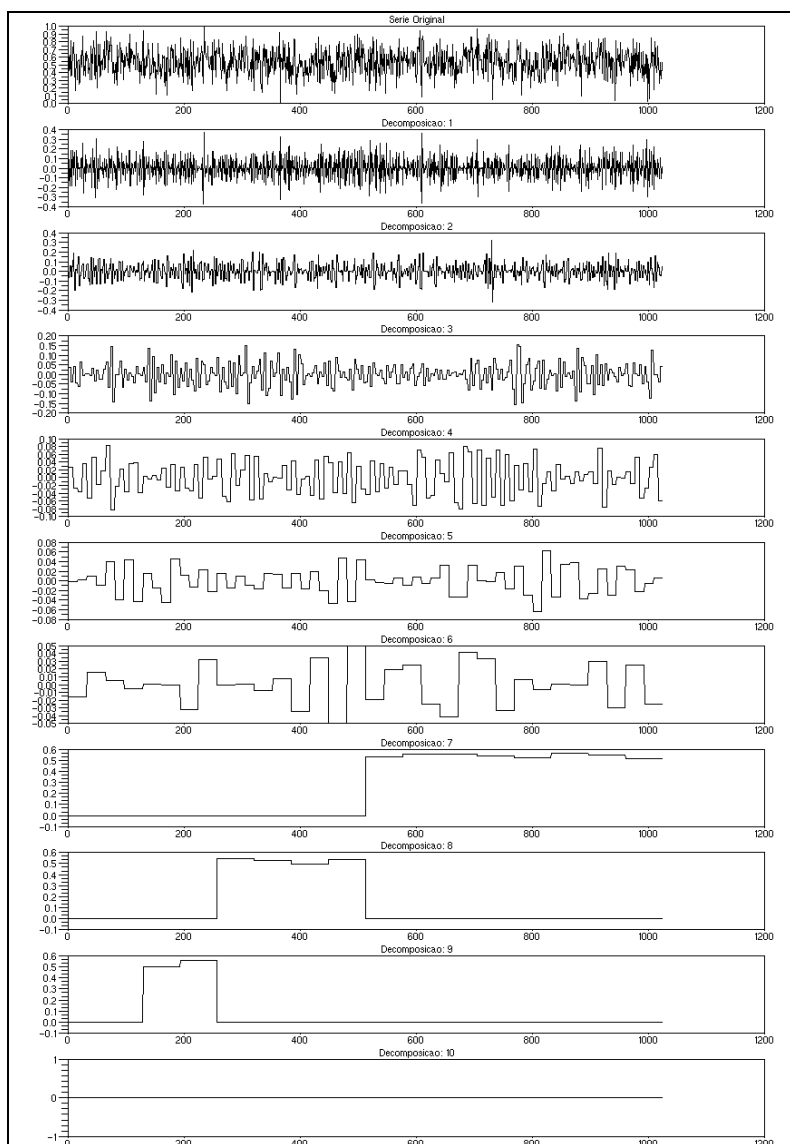


FIGURA 5.1 – Exemplos de decomposições da série gaussiana (topo) com a ondeleta discreta de Haar.

As diferentes resoluções de um sinal, em geral, possibilitam a caracterização das estruturas intrínsecas de variabilidade do mesmo. Isto é, através das escalas mais

grosseiras é possível identificar as características globais do sinal (baixas frequências) e, das escalas mais finas pode-se identificar os detalhes do mesmo (altas frequências). Dessa maneira, as decomposições com a ondeleta de Daubechies permitem representar as irregularidades temporais de cada componente espectral intrínseca de uma série curta evidenciando suas escalas de variabilidade no tempo, preservando o padrão da variabilidade do sinal decomposto analisado.

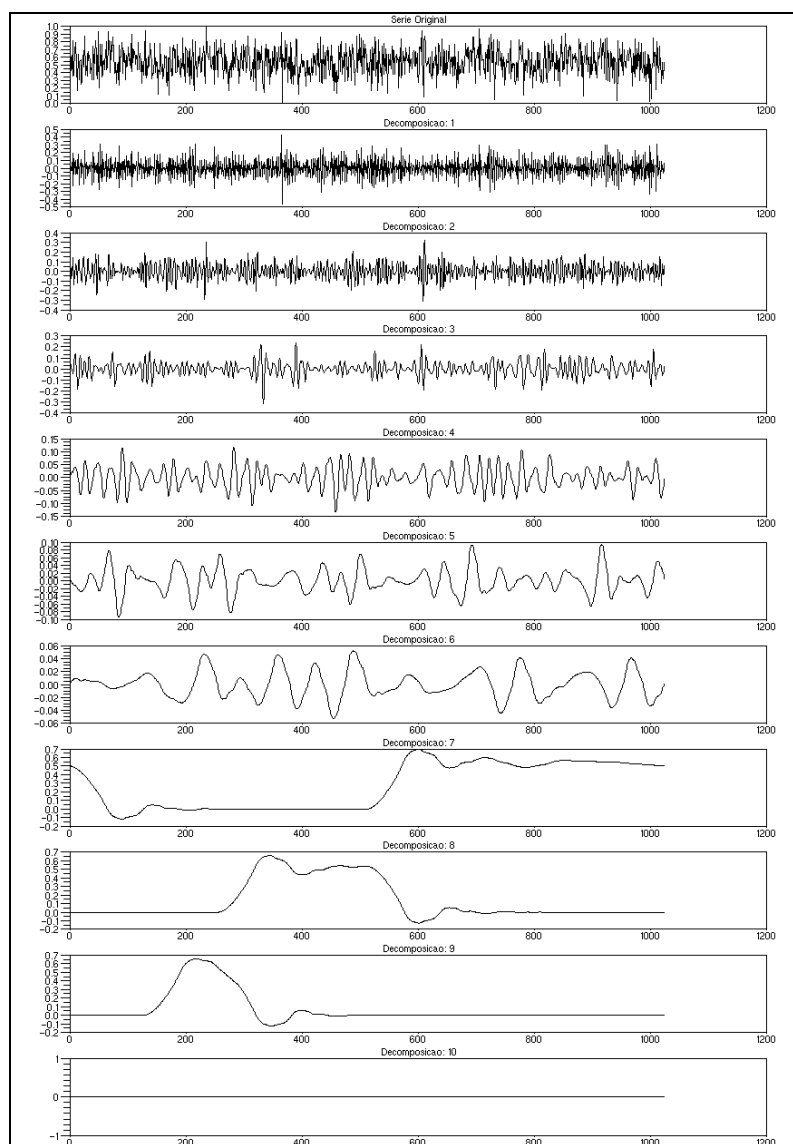


FIGURA 5.2 – Exemplos de decomposições da série gaussiana (topo) com a ondeleta discreta de Daubechies.

5.2 Cálculo das Assimetrias das Componentes Espectrais

O primeiro momento gradiente (g_1^a) da técnica de GPA1D, pode ser utilizado para caracterizar padrões de variabilidades irregulares no domínio temporal (Assireu et al., 2004). Por não se basear em parâmetros estatísticos no cálculo da assimetria do padrão global de uma ST curta (matriz 32×32 relativa a séries com 1024 pontos), a medida de assimetria, via GPA1D, apresenta a importante característica de ser menos dependente do comprimento da série, quando comparada aos demais métodos estatísticos.

O padrão de variabilidade local, de menor escala, pode ser obtido através das subamostras constituídas por 9 pontos (matrizes 3×3) obtidas ao longo da série por meio de uma janela corrida, isto é, com deslocamento ponto a ponto.

Na Figura 5.3 são apresentados três padrões elementares (constituídos por 9 pontos) que geram matrizes cujos campos gradientes são totalmente simétricos (vetores com a mesma orientação). Nestes casos, o grau de fragmentação assimétrica é nulo, resultando na remoção dos vetores com a mesma orientação.

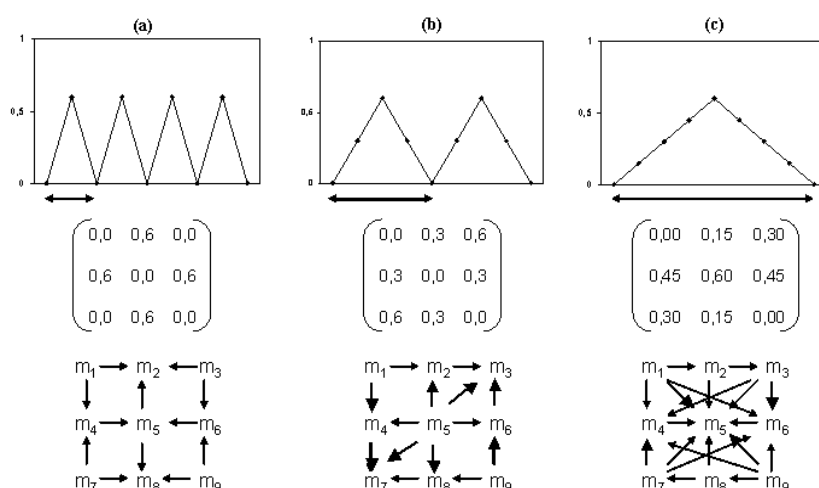


FIGURA 5.3 – Padrões de simétricos elementares (9 pontos) e suas respectivas matrizes e campos-gradiente.

Diferente do exemplo anterior, na Figura 5.4 são apresentados dois diferentes padrões assimétricos elementares, também constituídos por 9 pontos. Apesar de apresentarem campos gradientes assimétricos (vetores com orientações diferentes) com o mesmo número de vetores, o padrão fragmentado aleatoriamente (Figura 5.4-b) possui assimetria maior, que será quantificada pela conexão de pontos médios dos vetores com linhas retas, gerando a triangulação de Delaunay (Adelsbrunner, 1993).

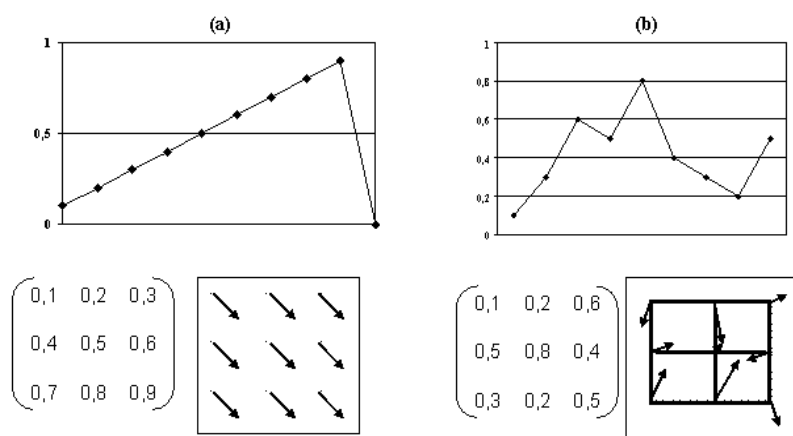


FIGURA 5.4 – Padrões de assimétricos elementares (9 pontos) e suas respectivas matrizes e campos-gradiente.

5.3 Espectro de Assimetrias

O espectro de assimetrias pode ser definido como uma relação entre as decomposições por ondeletas e a GPA1D, com a finalidade de caracterizar o padrão de variabilidade de uma ST.

O cálculo para obtenção do espectro de assimetrias será realizado através de duas abordagens. A metodologia utilizada, nas duas abordagens, considera ST curtas e não-estacionárias constituídas por 1024 pontos, normalizadas entre 0 e 1. As séries são decompostas nos primeiros 10 níveis de frequências, com auxílio da transformada de ondeletas discreta de Daubechies (Daubechies, 1988) (Figura 5.2). As decomposições

obtidas são armazenadas em arquivos independentes, com a finalidade de serem utilizadas como série de entrada na aplicação da GPA1D.

5.3.1 Primeira Abordagem

Os passos seguidos para o cálculo do espectro de assimetrias através da primeira abordagem são sintetizados pelo seguinte algoritmo.

1. Série temporal (1024 pontos normalizados)
2. Decomposições por ondeletas
 1. Soma das últimas 5 decomposições
3. GPA1D
 1. Janelas NxN
 2. Séries temporais de decomposições
 3. Valores de $\langle g_1^a \rangle$ locais
4. GPA1D
 1. Janela 32x32
 2. Valores de g_1^a globais
5. Relação
 1. Extensão Normalizada da Assimetria
6. Parâmetro quantitativo

FIGURA 5.5 – Algoritmo generalizado para obtenção do espectro de assimetrias (Primeira abordagem)

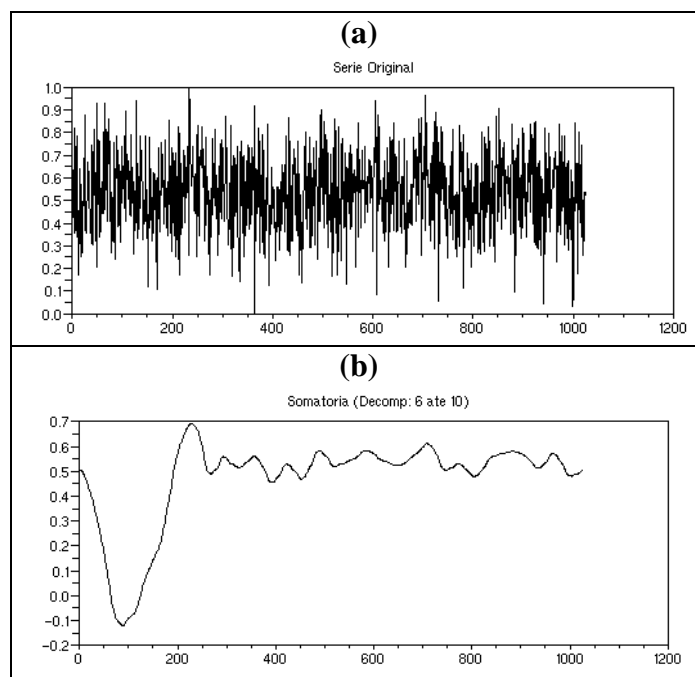


FIGURA 5.6 – (a) Série original; (b) Somatória das cinco últimas decomposições do sinal.

A primeira abordagem para cálculo do espectro de assimetrias considera uma ST curta (1024 pontos), em nosso caso a série gaussiana, e realiza 10 decomposições discretas com a ondeleta mãe de Daubechies na série original. As cinco primeiras decomposições são utilizadas, enquanto que, devido ao formato suavizado observado a partir da sexta decomposição (Figura 5.2), optou-se pela somatória das cinco últimas decomposições (Figura 5.6). Dessa maneira, nesta abordagem serão utilizadas 6 séries de decomposições.

A técnica de GPA1D foi aplicada nas séries de decomposições de duas maneiras. Ou seja, variando-se as dimensões das matrizes elementares de acordo com os níveis de decomposições. Isto é, para os primeiros níveis (relacionados às altas frequências) considera-se matrizes de dimensões menores e para os demais níveis (relacionados às baixas frequências) consideram-se matrizes de dimensões maiores. As dimensões das matrizes foram definidas da seguinte maneira: 3×3 para os 2 primeiros níveis, 4×4 para o terceiro, 8×8 para o quarto, 16×16 para o quinto e sexto níveis, sendo que este último representa a somatória dos últimos cinco níveis de decomposição. Em seguida, a

GPA1D foi novamente aplicada nas séries de decomposições considerando-se a dimensão fixa 32×32 .

Os resultados do $\langle g_1^a \rangle$ (locais) para as diferentes dimensões das matrizes elementares são apresentados na Tabela 5.1.

TABELA 5.1 – Resultados de $\langle g_1^a \rangle$ locais, obtidos após a aplicação da GPA1D nas séries de decomposições.

Decomposições	Matrizes Elementares	$\langle g_1^a \rangle$
1	3x3	1.222
2	3x3	1.111
3	4x4	1.375
4	8x8	1.781
5	16x16	1.926
6	16x16	1.930

Os resultados do g_1^a (globais) considerando (dimensão das matrizes elementares 32×32) são apresentados na Tabela 5.2.

TABELA 5.2 – Resultados de g_1^a globais, obtidos após a aplicação da GPA1D nas séries de decomposições.

Decomposições	g_1^a
1	1.978
2	1.981
3	1.981
4	1.977
5	1.968
6	1.986

Convencionamos caracterizar o padrão de variabilidade de uma ST curta através do seguinte parâmetro:

$$g_1^{a,R} = \left| \frac{R-r}{R} \right| \quad (5.1)$$

, onde R corresponde ao valor de g_1^a da matriz constituída por N pontos (escala global, em nosso caso 32×32) e r corresponde ao valor de $\langle g_1^a \rangle$ das matrizes cujas dimensões variam de acordo com a decomposição (escala local). Entre os diversos valores testados, o presente parâmetro denominado, provisoriamente, *extensão normalizada da assimetria* ($g_1^{a,R}$), é aquele que captura e diferencia com maior robustez a assinatura dos diferentes regimes de variabilidades irregulares (caótico, reativo-difusivo e turbulento).

TABELA 5.3 – Extensão normalizada da assimetria ($g_1^{a,R}$), calculada para a série gaussiana.

Decomposições	$g_1^{a,R}$
1	0.382
2	0.439
3	0.306
4	0.099
5	0.021
6	0.029

A partir destes resultados foi possível gerar um gráfico que representa o *Espectro de Assimetrias* (Figura 5.7). Neste gráfico, o eixo x representa as frequências da série, calculadas através da razão entre a frequência original e 2^n , onde n corresponde ao nível de decomposição (para o exemplo considerado, n varia entre 1 e 9). O eixo y do gráfico representa os valores da $g_1^{a,R}$ obtidos após a aplicação de 5.1.

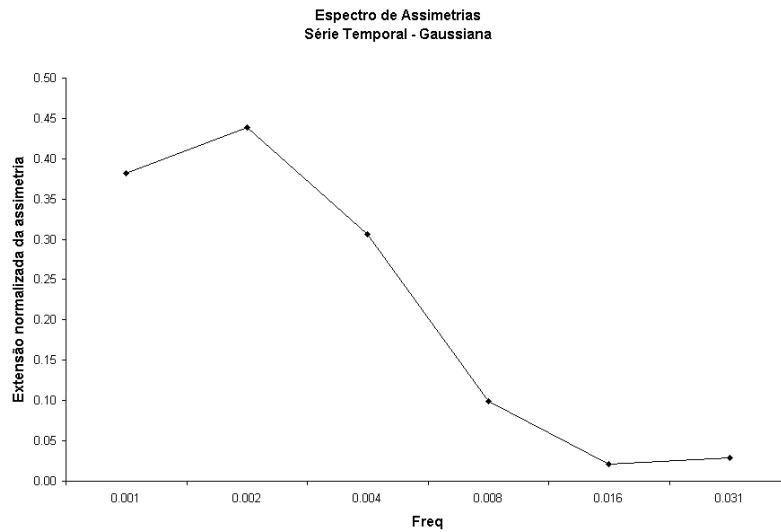


FIGURA 5.7 – Espectro de assimetrias para a série gaussiana.

A metodologia exposta anteriormente foi aplicada na ST canônica com padrão de variabilidade caótico 2D (Mapa de Hénon), obtendo-se os seguintes resultados (Tabela 5.4). O gráfico que representa este espectro é apresentado na Figura 5.8.

TABELA 5.4 – Extensão normalizada da assimetria ($g_1^{a,R}$) calculadas para a ST canônica com padrão de variabilidade caótico 2D (Mapa de Hénon).

Decomposições	$g_1^{a,R}$
1	0.382
2	0.440
3	0.306
4	0.102
5	0.022
6	0.026

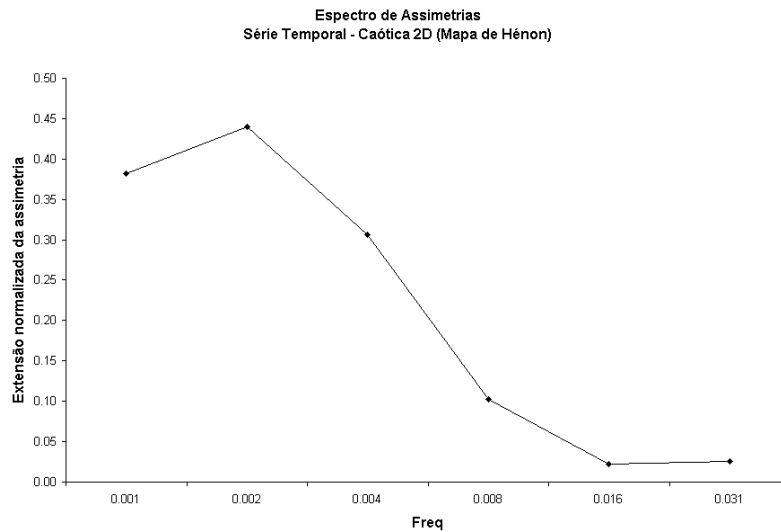


FIGURA 5.8 – Espectro de assimetria para dados canônicos com padrão de variabilidade caótico 2D (Mapa de Hénon).

A diferença entre os espectros de assimetrias (série gaussiana (Figura 5.7) e série caótica 2D (Figura 5.8)) foi calculada, com a finalidade de se obter um valor quantitativo para auxiliar a análise dos resultados.

Em teoria da probabilidade e teoria da informação, a divergência de Kullback-Leibler (Kullback e Leibler, 1951) é definida como uma distância natural de uma distribuição de probabilidade p para uma distribuição de probabilidade arbitrária q . Tipicamente, p representa os dados (observações) e q uma teoria, um modelo ou uma descrição aproximada de p . A distância de Kullback-Leibler é definida da seguinte maneira:

$$d_{KL}(p, q) = \sum_x p(x) \log \frac{p(x)}{q(x)} \quad (5.2)$$

, onde d_{KL} é a informação perdida quando q é usado para aproximar p . Em nosso caso, p será definida como uma série qualquer e q será a série de referência, isto é, a gaussiana.

A tabela abaixo (Tabela 5.5) corresponde aos resultados obtidos após a realização dos cálculos da distância de Kullback-Leibler entre os espectros de assimetria das séries gaussianas e aqueles obtidos para os demais padrões de variabilidade, apresentados no Capítulo 3.

TABELA 5.5 – Distâncias de Kullback-Leibler entre os espectros de assimetria (série gaussiana e dados canônicos).

Padrões de Variabilidades Canônicos	Distâncias
Caótico 1D (Mapa Logístico)	0.003 ± 0.002
Caótico 2D (Mapa de Hénon)	0.003 ± 0.002
Caótico 3D (Modelo de Lorenz)	0.006 ± 0.008
Reativo-Difusivo (Eq. de Kuramoto-Sivashinsky)	0.004 ± 0.007
Turbulento (Velocidade do vento – projeto LBA)	0.029 ± 0.002

5.3.2 Segunda Abordagem

Os passos seguidos para obtenção do espectro de assimetrias na segunda abordagem são apresentados através do seguinte algoritmo.

1. Série temporal (1024 pontos normalizados)
2. Decomposições por ondeletas
~~1. Soma das últimas 5 decomposições~~
3. GPA1D
 1. **Janelas 3x3**
 2. Séries temporais de decomposições
 3. Valores de $\langle g_1^a \rangle$ locais
4. GPA1D
 1. Janela 32x32
 2. Valores de g_1^a globais
5. Relação
 1. Extensão Normalizada da Assimetria
6. Parâmetro quantitativo

FIGURA 5.9 – Algoritmo generalizado para obtenção do espectro de assimetrias (Segunda abordagem)

Conforme observado no algoritmo acima (Figura 5.9) a metodologia original deve ser seguida, isto é, considerando-se uma ST curta (1024 pontos), em nosso caso a série gaussiana, e realizando-se 10 decomposições discretas com a ondeleta mãe de Daubechies.

Uma diferença em relação a primeira abordagem, consiste em considerar as nove primeiras séries de decomposições (Figura 5.9). Na segunda abordagem, somente a última série de decomposição será descartada, por apresentar sinal contínuo com valor zero (Figura 5.2). A segunda diferença corresponde a aplicação da técnica de GPA1D, ou seja, nesta abordagem deve-se considerar matrizes elementares de dimensões 3×3 (locais) em todos os níveis de decomposições. Como ocorreu na primeira abordagem, a GPA1D deve ser novamente aplicada nas séries de decomposições considerando matrizes de dimensões 32×32 (globais).

Os resultados de $\langle g_1^a \rangle$ locais e de g_1^a globais, referentes à ST gaussiana, obtidos para cada série de decomposição, são apresentados na Tabela 5.6.

TABELA 5.6 – Resultados de $\langle g_1^a \rangle$ locais e g_1^a globais, obtidos após a aplicação da GPA1D nas séries de decomposições.

Decomposições	$\langle g_1^a \rangle$	g_1^a
1	1.222	1.978
2	1.111	1.981
3	1.111	1.981
4	1.111	1.977
5	1.000	1.968
6	1.000	1.981
7	1.000	1.962
8	0.778	1.937
9	0.000	1.919

Os gráficos referentes ao espectro de assimetrias calculados para as séries gaussiana e canônica com padrão de variabilidade caótico 2D (Mapa de Hénon) são apresentados nas Figuras 5.10 e 5.11.

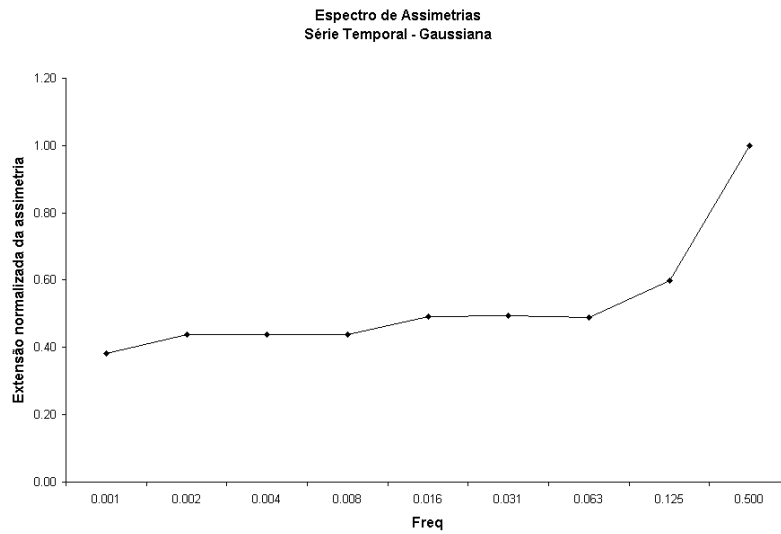


FIGURA 5.10 – Espectro de assimetrias para a série gaussiana.

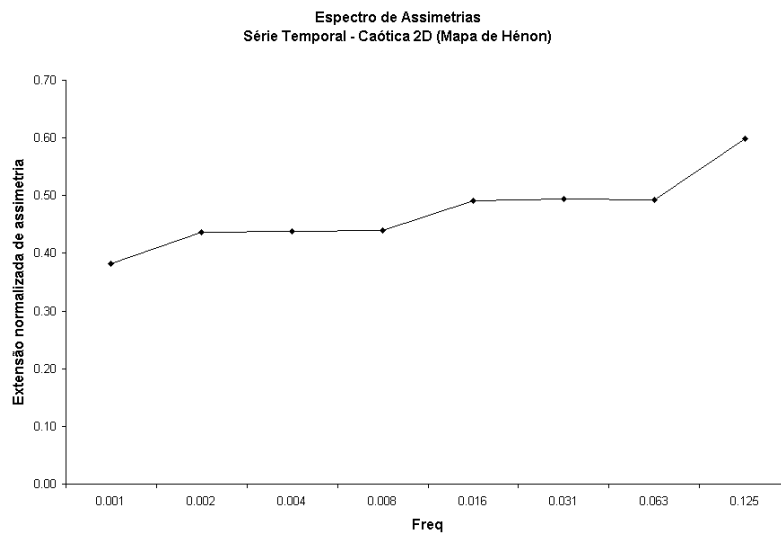


FIGURA 5.11 – Espectro de assimetria para dados canônicos com padrão de variabilidade caótico 2D (Mapa de Hénon).

A tabela abaixo (Tabela 5.7) corresponde aos resultados obtidos referentes à distância de Kullback-Leibler (5.2) entre os espectros de assimetria das séries gaussiana e aqueles obtidos para os demais padrões de variabilidade, apresentados no Capítulo 3.

TABELA 5.7 – Distâncias de Kullback-Leibler entre os espectros de assimetria (série gaussiana e dados canônicos).

Padrões de Variabilidades Canônicos	Distâncias
Caótico 1D (Mapa Logístico)	0.008 ± 0.011
Caótico 2D (Mapa de Hénon)	0.002 ± 0.001
Caótico 3D (Modelo de Lorenz)	0.009 ± 0.010
Reativo-Difusivo (Eq. de Kuramoto-Sivashinsky)	0.021 ± 0.012
Turbulento (Velocidade do vento – projeto LBA)	0.045 ± 0.012

5.4 Resultados Preliminares – Primeira e Segunda Abordagens

Com a aplicação das abordagens anteriormente definidas, nos conjuntos de dados de estudos de casos, foi possível obter o espectro de assimetrias das séries, cujos resultados são apresentados nas Tabelas 5.8 e 5.9.

TABELA 5.8 – Distâncias de Kullback-Leibler entre os espectros de assimetria (séries gaussiana e estudos de casos) – Primeira abordagem.

Estudos de casos	Distâncias
Temperatura do Ar (reservatório de Manso)	0.038
Temperatura do Ar (reservatório de Serra da Mesa)	0.025
Amplitudes de correntes oceânicas (componente Leste)	0.023
Amplitudes de correntes oceânicas (componente Norte)	0.041

TABELA 5.9 – Distâncias de Kullback-Leibler entre os espectros de assimetria (séries gaussiana e estudos de casos) – Segunda abordagem.

Estudos de casos	Distâncias
Temperatura do Ar (reservatório de Manso)	0.053
Temperatura do Ar (reservatório de Serra da Mesa)	0.026
Amplitudes de correntes oceânicas (componente Leste)	0.023
Amplitudes de correntes oceânicas (componente Norte)	0.038

Com a interpretação dos resultados apresentados nas Tabelas 5.8 e 5.9 pode-se concluir que, com base na metodologia dos espectros de assimetria, os padrões de variabilidade de temperatura do ar nos reservatórios são distintos (considerando-se as duas abordagens), apesar da semelhança na inspeção visual. Essa distinção sugere que,

possíveis modelos para os sistemas envolvidos devem levar em consideração aspectos não-lineares relacionados com as variáveis associadas a cada reservatório. Em relação ao dados de amplitudes de correntes oceânicas, concluí-se que as mesmas observações são pertinentes para as componente Leste e Norte.

Entretanto, do ponto de vista de validação fenomenológica dos resultados, recomenda-se à análise de outras séries equivalentes para os estudos de casos tomados como exemplo de aplicação.

CAPÍTULO 6

CONCLUSÃO

A partir da realização deste trabalho, pode-se concluir que as características oferecidas pelo ambiente de SCDP *Scilab* contemplam as principais funções e aplicações fornecidas pelos ambientes de softwares proprietários: MATLAB® e IDL. Entretanto, deve-se observar que, na fase inicial dos trabalhos, diversas pesquisas foram realizadas com a finalidade de simular algumas rotinas utilizadas pelos softwares proprietários, e ainda não implementadas no *Scilab*.

As técnicas estatísticas avançadas (transformada de ondeletas e GPA1D), implementadas em SCDP *Scilab*, foram aplicadas em diferentes conjuntos de dados (canônicos e ambientais). Os resultados obtidos foram comparados, considerando-se os ambientes de SCDP (*Scilab*) e de softwares proprietários (MATLAB® e IDL), e mostraram-se muito semelhantes para todos os casos analisados. Entretanto, devem-se destacar as limitações relacionadas às resoluções de cores, observadas nos resultados dos periodogramas em *Scilab*.

As implementações das técnicas de decomposições por ondeletas e de GPA1D em um ambiente de SCDP comum viabilizaram, a partir de sua utilização conjunta, os espectros de assimetria. A utilização dos espectros de assimetria pode complementar a caracterização de diferentes padrões de variabilidade em ST curtas. Os resultados obtidos mostraram-se satisfatórios, uma vez que conseguiram identificar, de maneira quantitativa, os valores para caracterização dos diferentes padrões de variabilidades apresentados, indicando na direção de um novo paradigma para análise de séries não-estacionárias, principalmente aquelas restritas em relação ao número de pontos.

A partir dos resultados obtidos durante a realização deste trabalho, espera-se estimular a realização de análises de ST curtas em ambiente de SCDP. Adicionalmente, a utilização de SCDP pode ser uma maneira eficiente de sumarizar as burocracias financeiras de instituições públicas e privadas de ensino e pesquisa, uma vez que não existem necessidades de compra e licenciamento dos produtos de softwares, e de viabilizar o desenvolvimento de pesquisas, com todos os benefícios decorrentes dessa filosofia.

6.1 Trabalhos Futuros

Novos testes de validação para obtenção do espectro de assimetrias: Sugere-se utilizar dados misturados aleatoriamente e outros padrões de variabilidade.

Ajuste de polinômio: Sugere-se ajustar um polinômio aos espectros de assimetrias e, através deste, gerar uma nova série com maior número de pontos para cálculo da distância de Kullback e Leibler.

Confecção de interface gráfica: Sugere-se a implementação de uma interface gráfica amigável para auxiliar os usuários nos processos de entrada de dados e escolha da técnica de análise.

Análises de ST via Internet: Sugere-se a implementação de uma interface amigável com acesso à Internet com a finalidade de viabilizar a realização de análises de ST, através de decomposições por ondeletas, GPA1D e espectro de assimetrias, para usuários geograficamente distribuídos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADELSBRUNNER, H. **Geometric algorithms**. North-Holland: Elsevier, 1993.
- ADDISON, P. The little wave with the big future. **Physics World**, v.17, n.3, p. 35-39, 2004.
- AGUIRRE, L.A. **Introdução à identificação de sistemas**. técnicas lineares e não-lineares aplicadas a sistemas reais. 2. ed. Belo Horizonte: UFMG, 2004.
- ALLIGOOD, K.T., SAUER, T.D., YORKE, J.A. **Chaos an introduction to dynamical systems**. New York: Spring-Verlag, 1997.
- ANASTASIADIS, A., Vlahos, L. Particle acceleration in an evolving active region by an ensemble of shock waves. **Astrophysics Journal**, n.428, p. 819-826, 1994.
- ASSIREU, A.T.; STECH, J.L.; NOVO, E.M.L.M.; LORENZZETTI, J.A., LIMA, I.B.T., CARVALHO, J.C. Aplicação do operador de fragmentação assimétrica (FA) na comparação de dados coletados in situ por diferentes sensores e transmitidos pelos satélites brasileiros SCD e CBERS: um exemplo de aplicação ao sistema de monitoramento ambiental (SIMA). In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12. (SBSR), 16-21 abr. 2005, Goiânia. **Anais...** São José dos Campos: INPE, 2005, p. 2455-2462. CD-ROM. ISBN 85-17-00018-8. Publicado como: INPE-12687-PRE/7977. Disponível em: <[http://marte.dpi.inpe.br/rep-
/ltid.inpe.br/sbsr/2004/11.26.18.43](http://marte.dpi.inpe.br/rep-
/ltid.inpe.br/sbsr/2004/11.26.18.43)>. Acesso em: 24 abr. 2007. rep: ltid.inpe.br/sbsr/2004/11.26.18.43.

ASSIREU, A.T., LORENZZETTI, J.A., NOVO, E.M.L.M., STECH, J.L., BRAGA, C.Z.F., LIMA, I.B.T. Aplicação do operador de fragmentação assimétrica (FA) na caracterização de controles geomorfológicos em reservatórios hidroelétricos. **Revista Brasileira de Geociências**, v.32, n.4, 2004.

ASSIREU, A.T., LORENZZETTI, J.A., ARAÚJO, C.E.S., ROSA, R.R. Distinção quantitativa de feições oceanográficas observadas nas imagens termais a partir do operador de fragmentação assimétrica: indicações preliminares. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 11., 5-10 abr. 2003, Belo Horizonte. **Anais...** São José dos Campos: INPE, 2003. p. 1485 - 1491. 1 CD-ROM. Publicado como: INPE--PRE/. Disponível em: <<http://marte.dpi.inpe.br/rep- /ltid.inpe.br/sbsr/2002/11.08.14.09>>. Acesso em: 24 abr. 2007. rep: ltid.inpe.br/sbsr/2002/11.08.14.09.

ASSIREU, A.T., ROSA, R.R., LORENZZETTI, J.A., VIJAYKUMAR, N.L., REMPEL, E.L., RAMOS, F. M., Sá, L.D. ABREU, BOLZAN, M.J.A., ZANANDREA, A. Gradient Pattern analysis of short nonstationary time séries: an application to lagrangian data from satellite tracked drifters. **Physica D: Nonlinear Phenomena**, v.168, p. 397-403, 2002.

BAK, P., TANG, C., WIESENFELD, K. Self-organized criticality. **Phys. Rev. Letters**, 59(4), p. 364-374, 1988.

BOLZAN, M.J.A. Análise da transformada em ondeletas aplicada o sinal geofísico. **Revista Brasileira de Ensino em Física**, v.26, n.1, p. 37-41, 2004.

BOLZAN, M.J.A. **Um estudo sobre características do subdomínio inercial da turbulência desenvolvida em escoamento acima e dentro da copa da floresta amazônica de terra firme e acima do pantanal matogrossense.** 2002-06. 164 p. (INPE-10394-TDI/920). Tese (Doutorado em Meteorologia) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2002. Disponível em: <[84](http://mtc-</p></div><div data-bbox=)

m15.sid.inpe.br/rep-/sid.inpe.br/jeferson/2004/02.13.09.22>. Acesso em: 24 abr. 2007.
rep: sid.inpe.br/jeferson/2004/02.13.09.22.

BOLZAN, M.J.A. Estudo da influência das estruturas coerentes e da rugosidade na estimativa de fluxos turbulentos sobre o pantanal. 1998-03-12. 138 p. (INPE-7500-TDI/715). Dissertação (Mestrado em Meteorologia) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 1998. Disponível em: <<http://MTC-m13.sid.inpe.br/rep-/sid.inpe.br/MTC-m13@80/2005/08.05.15.25>>. Acesso em: 24 abr. 2007. rep: sid.inpe.br/MTC-m13@80/2005/08.05.15.25.

BOX, G.E.P., JENKINS, G.M. Time series analysis: forecasting and control. San Francisco: Holden-Day, 1976.

CAMPANHARO, A.S.L.O. Análise de sinais turbulentos na copa da floresta amazônica: em busca de caos e estruturas coerentes. In: WORKSHOP DOS CURSOS DE COMPUTAÇÃO APLICADA DO INPE, 5. (WORCAP), 26 e 27 out. 2004, São José dos Campos. **Anais...** São José dos Campos: Instituto Nacional de Pesquisas Espaciais, 2004. CD-ROM, On-line. Disponível em: <<http://hermes2.dpi.inpe.br:1905/rep/dpi.inpe.br/hermes2@1905/2005/10.03.19.31>>. Acesso em: 24 abr. 2007. rep: dpi.inpe.br/hermes2@1905/2005/10.03.19.31.

CROSS, M. C., HOHENBERG, P.C. Pattern formation outside of equilibrium. Rev. Mod. Physics, v.65, p. 851-1112, 1993.

DAUBECHIES, I. Where the wavelets come from? – A personal point of view. Proceedings of the IEEE, v.84, n.4, p. 510-513, 1996.

DAUBECHIES, I. Orthonormal bases of compactly supported wavelets. Comm. Purê Appl. Math, v.41, p. 906-966, 1988.

DEVANEY, R.L. **A first course in chaotic dynamical systems**-theory and experiment. Massachusetts: Addison-Wesley Publishing Program, 1992.

DOMINGUES, M.O., MEDES JÚNIOR, O., COSTA, A.M. Algumas aplicações wavelet na análise de sinais atmosféricos. In: Congresso Temático de Dinâmica e Controle da SBMAC, 2003, São José dos Campos. **Anais...** São José dos Campos: ITA, 2003.

EMERY, W.J. THOMSON, R.E. **Data analysis methods in physical oceanography. Second and revised edition.** Amsterdam: Elsevier Science B.V., 2001.

FRISH, U. **Turbulence: the legacy of a. n. Kolmogorov.** Cambridge: Cambridge University Press, 1995.

FARGE, M. Wavelet transforms and their applications to turbulence. **Ann. Rev. Fluid Mech.**, v.24, p. 395-457, 1992.

GRAPS, A. An introduction to wavelets. **IEEE Computer Society**, v.2, n.2, p. 50-61, 1995.

GINZBURG, V.L., LANDAU, L.D. Zh. **Eksp. Teor. Fiz.**, 20, 1064, 1950.

HAAR, A. Zur theorie der orthogonalen funktionen-systeme. **Math. Ann.**, v.69, p. 331-371, 1910.

HÉNON, M. A two-dimensional mapping with a strange attractor. **Communications in Mathematical Physics**, v.50, p. 69-77, 1976.

KANEKO, K. Globally coupled chaos violates the law of large numbers but not the central-limit theorem. **Phys. Rev. Lett.**, 65, p. 1391-1394, 1990.

KANTELHARDT, J.W., ZSCHIEGNER, S.A., KOSCIELNY-BUNDE, E., BUNDE, A., HAVLIN, S., EUGENE STANLEY, H. Multifractal detrended fluctuation analysis of nonstationary time series. **Physica A: Statistical Mechanics and its Applications**, v316, Issues 1-4, p. 87-114, 2002.

KOLMOGOROV, A.N. Local structure of turbulence in an incompressible viscous fluid at very high Reynolds numbers. **Sov Phys Uspekhi.**, 10(6), p. 734-746, 1968.

KULLBACK, S., LEIBLER, R. A. On information and sufficiency. **Annals of Mathematical Statistics**, 22(1), p. 79-86, 1951.

KURAMOTO, Y., TSUZUKI, T. Persistent propagation of concentration waves in dissipative media far from thermal equilibrium. **Prog. Theor. Phys.**, 55, p. 356-369, 1976.

LILLY, J.M., PARK, J. Multiwavelet spectral and polarization analysis of seismic records. **Geophys. J. Int.**, v.122, p. 1001-1021, 1995.

LORENZ, E.N. Deterministic nonperiodic flow. **Journal of the Atmospheric Science**, 20, p. 130-141, 1963.

MORETTIN, P.A., TOLOI, C.M.C. **Análise de séries temporais**. São Paulo: Edgar Blücher, 2004.

MORLET, J. Sampling theory and wave propagation. **NATO ASI Series, Issues In Acoustic Signal/Image Processing and Recognition**, n.1, p. 233-261, 1983.

NORDEMANN, D.J.R. Periodicidades, tendências e previsão a partir de análise espectral dinâmica da série dos níveis do rio Paraguai em Ladário (1900/1995). **Pesquisa Agropecuária Brasileira**, v.33, Número Especial, p. 1787-1790, 1998.

PEITGEN, H.O., JÜRGENS, H., SAUPE, D. **Chaos and fractals. New frontiers of science.** New York: Springer-Verlag, 1992.

PERCIVAL, D.B., WALDEN, A.T. **Wavelets methods for time series analysis.** Cambridge: Cambridge University Press, 2000.

OSBORNE, A. R. and PROVENZALE, A. Finite correlation dimension for stochastic systems with power-law spectra. **Physica D: Nonlinear Phenomena**, v.35, Issue 3, p. 357-381, 1989.

PENG, C. K., BULDYREV, S.V., GOLDBERGER, A.L., HAVLIN, S., SCIORTINO, F., SIMONS, M., STANLEY, H.E. Long-range correlations in nucleotide sequences. **Nature**, 356, p. 168-171, 1992.

PIRES, P.S.M., ROGERS, D.A. Free/open source software: an alternative for engineering students. In: ASEE/IEEE Frontiers in Education Conference, 2002. Indianápolis. **Anais...** Indianápolis: ASEE/IEEE, 2002.

REMPEL, E.L., CHIAN, A.C.L., MACAU, E.E.N., ROSA, R.R. Analysis of chaotic saddles in high-dimensional dynamical systems: the Kuramoto–Sivashinsky equation. **Chaos**, v.14, n.3, p. 545-556, 2004.

RAMOS, F.M., BOLZAN, J.M.A., ABREU SÁ, L.D., ROSA, R.R. Atmospheric turbulence within and above an Amazon forest. **Physica D: Nonlinear Phenomena**, 193, p. 278-291, 2004.

ROSA, R.R., CAMPOS, M.R., RAMOS, F.R., VIJAYKUMAR, N.L., FUJIWARA, S., SATO, T. Gradient pattern analysis of structural dynamics: application to molecular system relaxation. **Braz. J. Phys.**, v.33, n.3, p. 605-610, 2003.

ROSA, R. R., PONTES, J., CHRISTOV, C. I., RAMOS, F. M., NETO, C. R., REMPEL, E. L., WALGRAEF, D. Gradient pattern analysis of Swift-Hohenberg dynamics: phase disorder characterization. **Physica A: Statistical Mechanics and its Applications**, v.283, Issue 1-2, p. 156-159, 2000.

ROSA, R.R. Padrões intermitentes em sistemas de tráfego contínuo em metrópoles. **Revista Unicsul**, n.5, p. 27-34, 1999.

ROSA, R.R., SHARMA, A.S., VALDIVIA, J.A. Characterization of asymmetric fragmentation patterns in spatially extended systems. **Int. J. Mod. Phys. C**, 10 (1), p. 147-163, 1999.

ROSA, R.R., SHARMA, A.S., VALDIVIA, J.A. Characterization of localized turbulence in plasma extended system. **Physica A: Statistical Mechanics and its Applications**, v.257, n.1, p. 509-514, 1998.

SAMSON, J.C. Pure states, polarized waves and principal components in the spectra of multiple geophysical time series. **Geophys. J. Roy. Astron. Soc.**, 72, p. 647-664, 1983.

SCHREIBER, T. Interdisciplinary application of nonlinear time series methods. **Phys. Rep.**, v.308, n.1, p. 1-64, 1998.

STECH, J.L., LIMA, I.B.T., NOVO, E.M.L.M., SILVA, C.M., ASSIREU, A.T., LORENZZETTI, J.A., CARVALHO, J.C., BARBOSA, C.C., ROSA, R.R. Telemetric monitoring system for meteorological and limnological data acquisition in aquatic environments. **Verh Internat. Verein Limnol.**, v.29, 2004.

STEVENSON, M.R.; LORENZZETTI, J.A.; STECH, J.L.; ARLINO, P.R.A. SIMA - an integrated environmental monitoring system. In: Simpósio Brasileiro de Sensoriamento Remoto, 1993, Curitiba. **Anais...** São José dos Campos: INPE, 1993.

SIVASHINSKY, G.I. Nonlinear analysis of hydrodynamics instability in laminar flows, I: Derivation of basic equations. **Acta Astron.**, v.4, p. 1177-1206, 1977.

SWINNEY, H.L., TSALLIS, C. Anomalous distributions, nonlinear dynamics, and nonextensivity. **Physica D: Nonlinear Phenomena**, v.193, Issue 1-4, p. 1-2, 2004.

THOMSON, D.J. Spectrum estimation and harmonic analysis. **Proc. IEEE**, v.70, n.9, p. 1055-1096, 1982.

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.

TRENBERTH, K.E. Spatial and temporal variations of the southern oscillation. **Quart. J. Roy. Meteor. Soc.**, n.102, p. 639-653, 1976.

VERHULST, P.F. Recherches mathématiques sur la loi d'accroissement de la population. **Nouv. mém. de l'Académie Royale des Sci. et Belles-Lettres de Bruxelles**, 18, p. 1-41, 1845.

VIEIRA, S. **Introdução à bioestatística. 5. ed.** Local: Editora Campus, 195 p., 1980.

ZANANDREA, A., NETO, C.R., ROSA, R.R., RAMOS, F.M. Analysis of geomagnetic pulsations using multiwavelets spectral and polarization method. **Physica A: Statistical Mechanics and its Applications**, v.283, Issues 1-2, p. 175-180, 2000.

ZEBENDEA, G.F., PEREIRA, M.G., NOGUEIRA Jr., E., MORET, M.A. Universal persistence in astrophysical sources. **Physica A: Statistical and Theoretical Physics**, v.349, Issues 3-4, p. 452-458, 2005.

WANG, B., WANG, Y. Temporal structure of the southern oscillation as revealed by waveform and wavelet analysis. **J. Climate**, v.9, Issue 7, p. 1586-1598, 1996.

APÊNDICE A

TEORIA DE ONDELETAS

A teoria de ondeletas é relativamente nova e têm sido continuamente desenvolvida nos últimos anos. Ondeletas são definidas como funções que satisfazem certas exigências matemáticas tais como, análises harmônicas em diferentes escalas e resoluções e de conjuntos de funções ortogonais e não-ortogonais, e podem ser empregadas na representação de dados ou de outras funções.

O termo ondeleta foi primeiramente citado no início do século, em 1910, por Alfred Haar com a introdução de um sistema completo de funções ortogonais. No entanto, o conceito de ondeletas foi formalizado no início da década de oitenta, a partir de artigos sobre a análise de sinais sísmicos em geofísica, publicados por (Morlet, 1983).

A teoria de ondeletas é baseada na análise de dados a partir de um conjunto de funções bases. No entanto, a utilização de funções para analisar de dados foi primeiramente proposta por Joseph Fourier, no início do século XIX. Com sua teoria de análise de freqüências, Fourier afirmou que qualquer função periódica $f(x)$ poderia ser representada como uma somatória de senos e co-senos. Entretanto, os métodos de Fourier estimam sinais em termos de suas componentes de freqüência e amplitude, desconsiderando as informações relacionadas ao tempo. Esta limitação de localização no tempo torna a teoria de Fourier inadequada para aplicações em sinais não-estacionários, que são caracterizados por mudanças locais de freqüências.

Diante desse contexto, muito esforço científico foi empregado na descoberta de funções apropriadas, sem considerar senos e co-senos, para a aproximação de sinais. O resultado corresponde a um conjunto de funções irregulares contidas em domínio finito, que pode ser muito bem empregado na aproximação de dados salientes e descontínuos, ou seja, as

ondeletas. Trata-se da técnica mais apropriada para análises de sinais não-estacionários (Graps, 1995; Addison, 2004; Bolzan, 2004), devido a marcante característica que captura informações em diferentes escalas, para a realização de estudos em diferentes resoluções.

Grande dessa teoria foi desenvolvida de maneira independente em diferentes campos da ciência como, por exemplo, matemática, física, ciência da computação e engenharia. Durante os últimos anos, contribuições multidisciplinares têm viabilizado um elevado número de aplicações em diferentes áreas, destacando-se a análise de dados, cujo crescimento exponencial representa uma síntese de idéias antigas aliadas a resultados matemáticos recentes, a eficientes algoritmos computacionais e ao interesse da ampla comunidade (Daubechies et al., 1992).

Formalmente, para que uma função seja denominada ondeleta (ψ), a seguinte condição de admissibilidade deve ser satisfeita:

- A integral da função ondeleta deve ser zero, ou seja,

$$\int_{-\infty}^{+\infty} \psi(t) dt = 0. \quad (\text{A.1})$$

Isso garante que a função tenha forma semelhante à de uma onda. Além disso, a função ondeleta deve ter energia unitária, ou seja,

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

Isso garante que a função possua suporte compacto, ou com decaimento rápido de amplitude, garantindo a localização espacial (Domingues et al., 2003).

O termo “função ondeleta” é utilizado genericamente para se referir a funções ondeletas ortogonal e não-ortogonal. No entanto, o termo “função base” refere-se somente ao conjunto de funções ortogonais. A utilização da base ortogonal implica no uso da transformada discreta de ondeleta, enquanto que, ondeletas não-ortogonais podem ser utilizadas tanto para transformadas discretas como para contínuas (Graps, 1995).

Para explicar uma função base, pode-se deixar, momentaneamente, o domínio das funções e recorrer ao domínio dos vetores. Ou seja, todo vetor bi-dimensional (x, y) é uma combinação dos vetores $(1, 0)$ e $(0, 1)$. Estes dois vetores são as bases para (x, y) , uma vez que o resultado da multiplicação de x por $(1,0)$ é o vetor $(x,0)$ e, de modo similar, multiplicando y por $(0, 1)$, obtêm-se $(0, y)$. Então, a soma dos vetores é (x, y) .

Adicionalmente, a valiosa propriedade de que os vetores bases são perpendiculares, ou ortogonais, entre si é satisfeita. Estes conceitos também se aplicam às funções bases. Ou seja, ao invés de vetores (x,y) temos uma função $f(x)$. Se $f(x)$ for um tom musical e A um conjunto qualquer de tons, então, poderíamos construir A pela adição de senos e cossenos utilizando combinações de amplitudes e frequências. Neste exemplo, seno e cosseno são as funções bases e serão ortogonais, mediante uma combinação de termos apropriada, tal que o produto interno seja igual a zero (Graps, 1995).

Uma família generalizada de ondeletas pode ser descrita da seguinte maneira,

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right). \quad (\text{A.3})$$

O parâmetro a se refere à escala, b é o parâmetro de translação ou localização da função ondeleta ($\psi_{a,b}$), geralmente, chamada de função ondeleta mãe. A variação do parâmetro a tem o efeito de dilatação (quando $a > 1$) e de contração (quando $a < 1$) na função ondeleta mãe (Figura A.1). Com isto, podem-se analisar, respectivamente, aspectos de longos e curtos períodos na série. À medida que b varia, o sinal é analisado localmente em torno desse ponto.

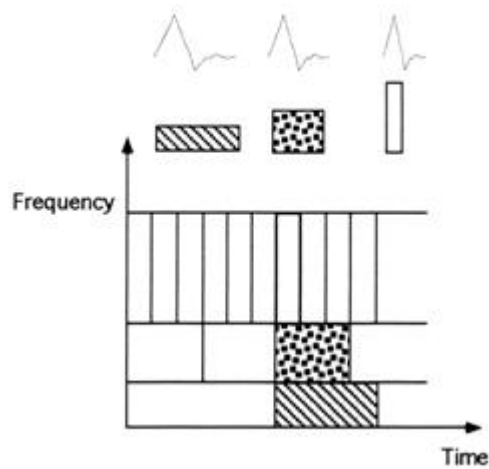


FIGURA A.1 – Variações proporcionais dos intervalos de tempo e de frequência.

FONTE: Graps (1995)

Funções ondeletas são classificadas como discreta ou contínua. A função ondeleta discreta mais simples é a de (Haar, 1910) (Figura A.2), utilizada na decomposição e filtragem de qualquer série temporal. Por ser discreta, essa categoria de funções trabalha com sinais temporais de comprimento, da ordem de potência de dois mais próxima, ou seja, $N = 2^d$, onde N representa o comprimento da série e d representa o número de frequências possíveis para a decomposição. O procedimento de decomposição do sinal em várias escalas possibilita o estudo de características particulares em cada frequência. Nesse sentido, é possível estudar fenômenos que ocorrem somente em determinadas frequências (Bolzan, 2004). Entretanto, devido à forma descontínua da onda, considera-se pobre o resultado correspondente à localização de frequência produzida, sendo esta função, em geral, empregada no entendimento básico da análise de ondeletas.

$$\psi(t) = \begin{cases} 1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{caso contrário} \end{cases} \quad (\text{A.4})$$

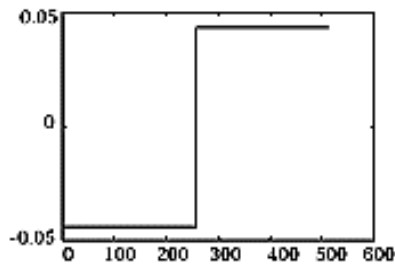


FIGURA A.2 – Função ondaleta “Haar”.

FONTE: Graps (1995)

Além de produzir espectros compostos por blocos discretos, esse tipo de função, quando aplicada na análise de séries temporais não-estacionárias, produz diferentes espectros. É possível construir funções ondaletas ortogonais utilizando uma ferramenta matemática conhecida como multiresolução (Domingues, 2003), onde a função mãe é obtida a partir de uma função de escalonamento. Com isso, bases com suporte compacto e grau de suavidade arbitrário podem ser construídas. Um exemplo desse tipo de construção são as funções de (Daubechies, 1988) (Figura A.3). Em geral, essas funções não possuem expressões analíticas e são não-simétricas. Para se obter funções simétricas, foram desenvolvidos procedimentos utilizando duas análises de multiresolução, o que permite a construção de famílias Bi-ortogonal de Daubechies.

Algumas funções ondaletas possuem estrutura fractal e a família das Daubechies é uma delas, como se pode observar na parte em destaque da Figura A.3 (Graps, 1995).

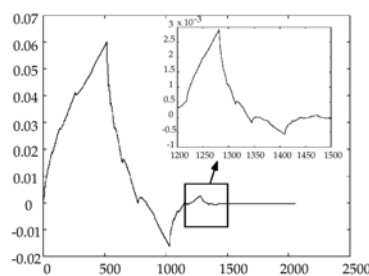


FIGURA A.3 – Função ondaleta “Daubechies”.

FONTE: Graps (1995)

Numa ondeleta contínua, os parâmetros de escala e localização assumem valores contínuos. Em geral, estas funções são representadas a partir de diagramas tridimensionais, com a finalidade de representar as relações existentes entre as componentes de diferentes frequências em função da escala temporal do sinal. Entretanto, para o cálculo de uma função ondeleta contínua, muitas das informações apresentadas referentes à escala ou tempos próximos são redundantes e isso faz com que este tipo de função possua um alto custo computacional. Uma das funções contínuas mais comum é a de (Morlet, 1983) (Figura A.4), a qual consiste de uma onda plana modulada por uma Gaussiana.

$$\psi(t) = \pi^{-1/4} \left(e^{i\xi t} - e^{-\xi^2/2} \right) e^{-t^2/2} \quad (\text{A.5})$$

, onde ξ representa uma frequência não-dimensional. Para satisfazer a condição de admissibilidade, em geral, escolhe-se $\xi = 5$ (Graps, 1995), tal que o maior e menor valores de ψ seja aproximadamente $1/2$. A ondeleta de Morlet é largamente utilizada na análise das frequências pertencentes ao sinal e na identificação de relações entre tempo e frequência (Bolzan, 2004). Estas funções são utilizadas na análise de séries temporais, onde suavidade e variações contínuas na amplitude da ondeleta são esperadas (Torrence e Compo, 1998).

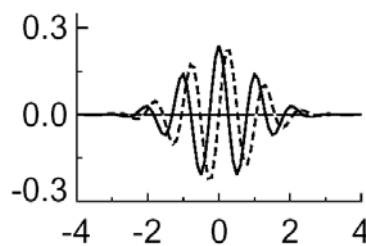


FIGURA A.4 – Função ondeleta “Morlet”. Parte real: linha cheia. Parte imaginária: linha pontilhada.

FONTE: Torrence e Compo (1998)

A ondeleta mãe conhecida como Chapéu Mexicano é a derivada de segunda ordem da função de densidade de probabilidade Gaussiana.

$$\psi(t) = \frac{2}{\sqrt{3\pi^{1/4}}} e^{-t^2/2} (1-t^2). \quad (\text{A.6})$$

Essa função possui suporte infinito, entretanto, seu suporte efetivo está no intervalo $[-5,5]$ (Figura A.5).

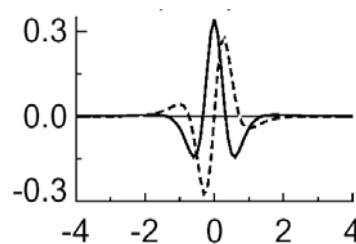


FIGURA A.5 – Função ondaleta “Chapéu Mexicano”. Parte real: linha cheia. Parte imaginária: linha pontilhada.

FONTE: Torrence e Compo (1998)

Funções ondaletas complexas tais como Morlet e Chapéu Mexicano, retornam informações sobre amplitude e fase e são melhores adaptadas para capturar comportamento oscilatório.

Um dos fatores críticos na análise de ondaletas é a escolha arbitrária da função mãe, $\psi_0(t)$. Ainda não existe na literatura um consenso de qual a melhor função ondaleta a ser utilizada. Comumente se aceita que a função ondaleta a ser utilizada possua um formato característico próximo das características encontradas na série temporal. Entretanto, alguns outros fatores podem ser considerados (Torrence e Compo, 1998) como, por exemplo:

- Velocidade de computação do produto interno com outras funções bases;
- Velocidade de superposição dessas funções;
- Boa localização espacial, de maneira a facilitar a identificação da posição do sinal;

- Boa frequência de localização, de maneira a facilitar a identificação da oscilação do sinal;
- Independência, de maneira que poucos elementos combinem com a mesma porção do sinal.

Além disso, pesquisadores procuram uma base em que os coeficientes, quando rearranjados em ordem decrescente, decrescem o mais rapidamente possível. Para medir tais taxas, ferramentas clássicas da análise harmônica são utilizadas, incluindo o cálculo do custo da função de informação. Exemplos destas funções incluem entropia, logaritmo da energia, cálculos de Gaus-Markov e dimensão teórica de uma seqüência (Graps, 1995).

APÊNDICE B

ANÁLISE DE PADRÕES-GRADIENTE

A técnica para análise de padrões-gradiente (GPA) baseia-se na aplicação de operadores computacionais que quantificam quebras de simetria e desordem de fase em escalas locais e globais de um dado padrão espaço-temporal. A GPA analisa padrões espacialmente gerados em duas dimensões (x, y) e é representado pela matriz das amplitudes:

$$M = L^{lxl} \{M(1, 1), \dots, M(i, j), \dots, M(l, l) \mid i, j \in I \text{ e } M \in \mathcal{R}\}$$

Essencialmente é uma grade quadrada (L), com as dimensões espaciais (x, y), discretizadas em lxl pontos, com $i = 1, \dots, l$ e $j = 1, \dots, l$. Assim, uma sucessão dinâmica de N grades, L_0, L_1, \dots, L_N está relacionada à visualização da evolução temporal de um envelope de amplitude $M_{x,y,t} \equiv \in (x; y; t)$.

Dentro do formalismo da GPA, um campo vetorial gradiente $G_t = \nabla[\in (x, y)]_t$, composto por V vetores r , onde cada vetor $r_{i,j}$, localizado na posição (i, j) do campo gradiente, caracterizado por sua norma e fase ($r_{i,j} = (r, \emptyset)$), pode ser descrito pela composição de quatro momentos gradientes: o momento gradiente de primeira ordem, g_1 , uma medida global do padrão da distribuição espacial de todos os vetores, isto é, de todas as normas e suas respectivas fases localizadas no campo gradiente; o momento gradiente de segunda ordem, g_2 , que determina a diversidade existente na distribuição dos módulos dos vetores, o momento gradiente de terceira ordem, g_3 , que determina a diversidade existente na distribuição das fases dos vetores e o quarto momento, g_4 , que determina a entropia do campo gradiente com base na distribuição dos números complexos associados a cada vetor. Esses momentos estão apresentados graficamente na Figura B.1.

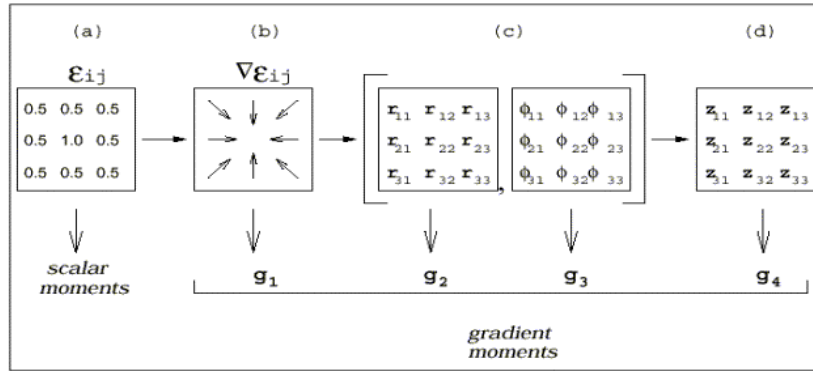


FIGURA B.1 – Momentos Gradientes: (a) Momento escalar, (b) Momento gradiente de primeira ordem, (c) Momentos gradiente de diversidades de módulo e fase e (d) Momento da medida global da norma e da fase.

Padrões bidimensionais sem fragmentação ou padrões fragmentados com simetria total possuem g_1^a nulo (não há irregularidades na distribuição). Este operador computacional gera uma medida da assimetria vetorial global do padrão espaço-temporal, quantificando, no tempo, as possíveis quebras de simetria. O primeiro momento (g_1^a) é definido por:

$$g_1^a = (I - L); (I \geq L > 0) \quad \text{B.1}$$

, onde L é o número de vetores assimétricos e I o número de barras de correlação gerado pela triangulação de Delaunay (Adelsbrunner, 1993) tomando um ponto local de cada vetor assimétrico como vértice.

A triangulação de Delaunay $T_D(L, I)$, neste contexto, é um campo fracionário com dimensão menor que a dimensão da grade (igual a 2). Quando não há correlação assimétrica no padrão, o número total de vetores assimétricos é zero, e então, por definição, o g_1^a é nulo. Para uma determinada grade de tamanho lxl , o padrão totalmente desordenado gera o mais alto valor para g_1^a , padrões conectivos compostos

por estruturas em forma de rolos e labirintos, por exemplo, apresentam valores específicos não-nulos para g_1^a .

B.1 Análise de Padrões-Gradiente de Séries Temporais (GPA1D)

A análise de ST é um procedimento estatístico em que os dados das séries são considerados como parte de um processo estocástico. As propriedades estatísticas de ST podem ser usadas para determinar as similaridades e as diferenças entre grupos de dados distintos. Com a finalidade de explorar uma maneira mais eficiente de caracterizar padrões de variabilidade em ST curtas e não-estacionárias, tem sido utilizado um operador do GPA1D.

Para aplicar a GPA1D às ST é necessária uma etapa preliminar, que consiste em converter uma seqüência de pontos de dados da série em uma seqüência de padrões-gradiente.

Dados experimentais consistem tipicamente em medidas discretas de um único observador $X(t) = \{X(t_1), \dots, X(t_i), \dots, X(t_n)\}$, cuja definição de tempo é dada por $t = t_{i+1} - t_i$. Para uma notação mais conveniente, escreve-se a série de tempo discreta como:

$$X(t) = \{X_1, \dots, X_i, \dots, X_n\}; X_i \in \mathcal{R}.$$

Para aplicar a GPA1D converte-se $X(t)$ em uma seqüência de $N \times N$ matrizes, com a condição $N \geq 3$. O comprimento da variabilidade temporal é dado por $\tau = (N - 1) \cdot t$. Note que a definição temporal mais elevada da variabilidade é obtida agrupando os dados em uma fileira de 3×3 matrizes. Este caso elementar, chamamos (9-1) padrão de variabilidade. Assim, cada seqüência de nove pontos da série de tempo dada produz a matriz M_{x_9} :

$$\begin{pmatrix} x_{11}^i & x_{12}^{i+1} & x_{13}^{i+2} \\ x_{21}^{i+3} & x_{22}^{i+4} & x_{23}^{i+5} \\ x_{31}^{i+6} & x_{32}^{i+7} & x_{33}^{i+8} \end{pmatrix}$$

O campo gradiente (∇M_{x_9}) é composto por nove vetores e especifica quantitativamente as variações em um dado ponto. É o indicador quantitativo apropriado das mudanças na distribuição destes máximos, ou seja, um indicador do padrão da variabilidade. Nesta aproximação, os valores relativos em vez dos valores absolutos da amplitude na matriz são dinamicamente relevantes.

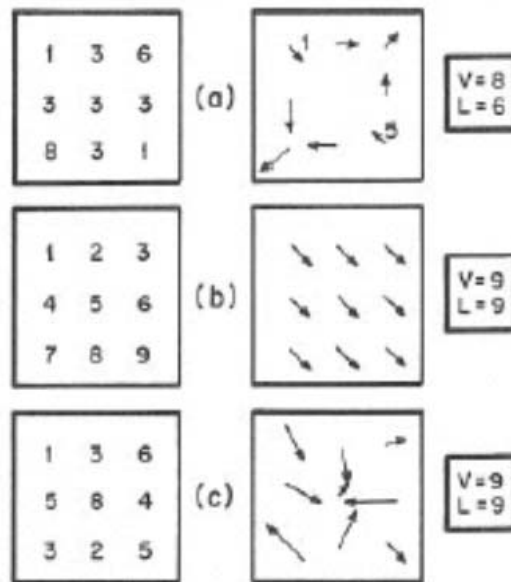


FIGURA B.2 – Matrizes representativas de campos vetoriais assimétricos. (a) 6 vetores assimétricos, (b) e (c) todos os vetores são assimétricos.

A distribuição espacial dos vetores no ∇M_{x_9} deverá conter muitos vetores com a mesma magnitude, com uma pequena tolerância, e estas formarão pares simétricos se tiverem orientações opostas, caso contrário, formarão pares assimétricos. Após remover todos os pares de vetores simétricos, o número de vetores assimétricos são denominados L . Qualquer quebra de simetria pode ser quantificada conectando-se os pontos médios

dos vetores L com as linhas retas I , gerando o campo de triangulação de Delaunay T_D (L, I).

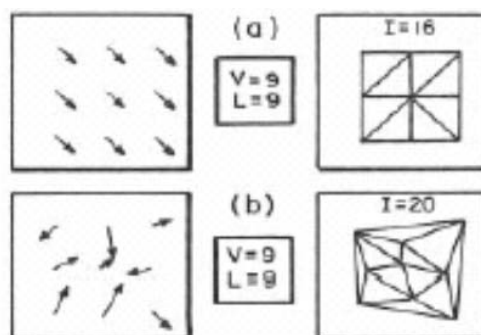


FIGURA B.3 – Campo de triangulação para (b) e (c) da Figura B.2.

Como esta técnica foi originalmente desenvolvida para tratar campos bidimensionais representados através de matrizes numéricas. Com o intuito de estender as aplicações do método para séries unidimensionais, foi feita a seguinte adaptação: cada nove pontos da série são convertidos em uma matriz 3×3 e o método é, desta forma, aplicado ao longo da série. Um valor médio, considerando cada um dos g_1^a obtidos ao longo da série é gerado (Figura B.4). Um exemplo de uma série unidimensional com nove pontos (Figura B.4-a) que é convertida em uma matriz 3×3 (Figura B.4-b) e finalmente os campos de gradiente e de triangulação são gerados (Figuras B.4-c e B.4-d). A escolha de se utilizar matrizes 3×3 deve-se à possibilidade das análises serem feitas na melhor janela de resolução e sem comprometimento da robustez do método.

Os únicos valores possíveis para o parâmetro g_1^a , calculados de ∇M_{x_y} , são restringidos ao domínio discreto D_9 composto por: $\{0.778; 0.889; 1.000; 1.111; 1.222; 1.333\}$, o que pode ser provado por considerações geométricas. Uma característica importante do primeiro momento gradiente é a sua alta sensibilidade para detectar pequenas modificações na variabilidade de um sinal que, em geral, não são detectadas por outras técnicas usuais.

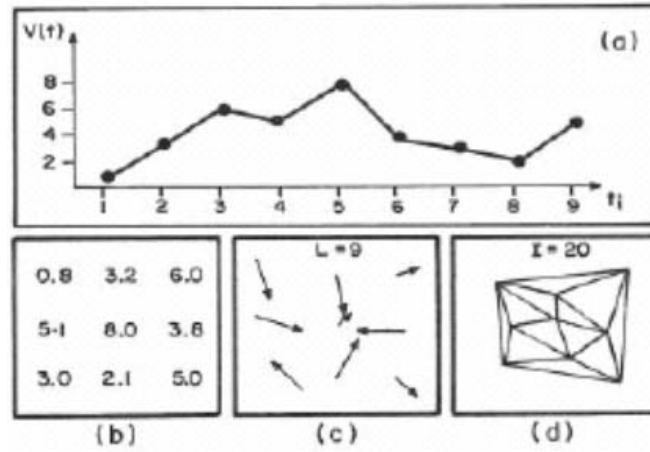


FIGURA B.4 – (a) Típico padrão de variabilidade, (b) Matriz correspondente, (c) Campo gradiente correspondente (9 vetores assimétricos), (d) Campo de triangulação com 20 linhas conectoras.

APÊNDICE C

SCILAB NA INTERNET

Desde 1994 o *Scilab* é distribuído livremente através da Internet, seguindo o conceito de software de domínio público. Através de uma rápida busca na seção *Downloads* (à esquerda da página) do sítio do *Scilab* na Internet (www.scilab.org) (Figura C.1) pode-se obter o software para diferentes sistemas operacionais como, por exemplo, Linux, UNIX e Windows (Figura C.2).

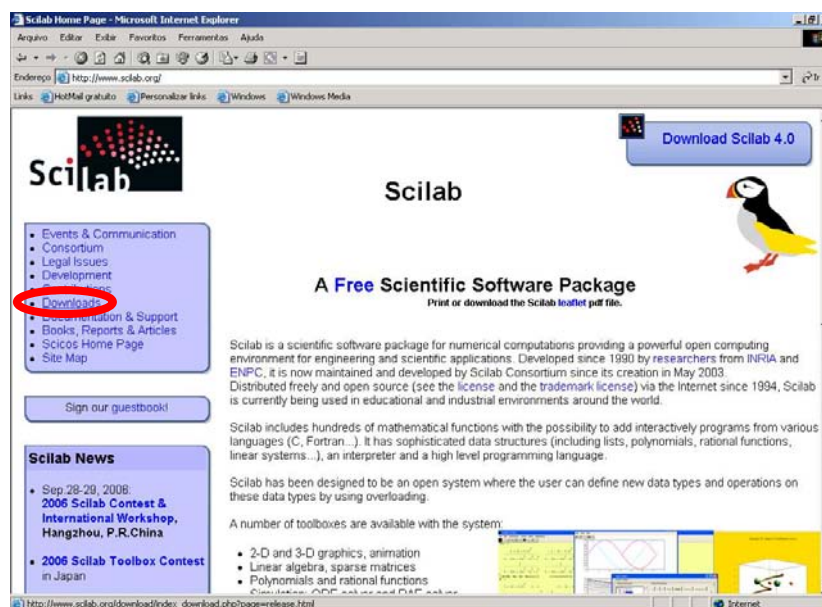


FIGURA C.1 – Página principal do sítio do *Scilab* na Internet. Destaque para a seção *Downloads* à esquerda da página.

Ao acessar a seção *Downloads* é possível aos usuários escolher a versão mais adequada para o sistema operacional em uso. A versão 3.1.1, pré-compilada para sistema operacional Windows 2000, foi a utilizada durante o desenvolvimento deste trabalho. Atualmente, versão 4.0 é considerada estável e recomendada para instalações pelo sítio do *Scilab* (Figura C.3).



FIGURA C.2 – Seção *Downloads* encontrada no sítio do *Scilab*. Destaque para a versão desenvolvida para o sistema operacional Windows.

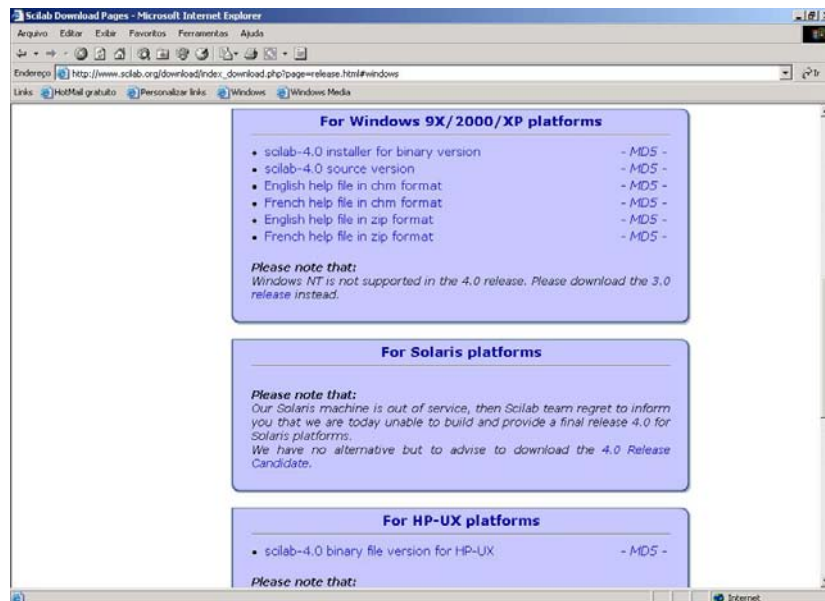


FIGURA C.3 – Opções de *Downloads* da versão 4.0 do *Scilab*, recomendada para instalações pelo sítio do software na Internet.

C.1 Algoritmos Desenvolvidos em *Scilab*

Os algoritmos referentes às técnicas de transformada de ondeletas, GPA e espectro de assimetrias, implementados em SCDP *Scilab*, durante o desenvolvimento deste trabalho, podem ser acessados livremente no CD-ROM com a versão final desta dissertação de mestrado.

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 programa 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). Constam 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. São aceitos tanto programas fonte quanto executáveis.

Pré-publicações (PRE)

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