INPE-16637-RPQ/842

ESTUDO EXPLORATÓRIO DO PADRÃO DE VARIABILIDADE DO CICLO DE CHEIA DO RIO PARAGUAI

Murilo da Silva Dantas

Relatório final da disciplina Princípios e Aplicações de Mineração de Dados (CAP-359) do Programa de Pós-Graduação em Computação Aplicada, ministrada pelo professor Rafael Santos.

Registro do documento original:
<http://urlib.net/sid.inpe.br/mtc-m19@80/2009/10.30.13.02>

INPE
São José dos Campos
2009

PUBLICADO POR:

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) 3945-6911/6923
Fax: (012) 3945-6919
E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO:
Presidente:
Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)
Membros:
Drª Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação
Dr. Haroldo Fraga de Campos Velho - Centro de Tecnologias Especiais (CTE)
Drª Inez Staciarini Batista - Coordenação Ciências Espaciais e Atmosféricas (CEA)
Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)
Dr. Ralf Gielow - Centro de Previsão de Tempo e Estudos Climáticos (CPT)
Dr. Wilson Yamaguti - Coordenação Engenharia e Tecnologia Espacial (ETE)
BIBLIOTECA DIGITAL:
Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)
Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)
Jefferson Andrade Ancelmo - Serviço de Informação e Documentação (SID)
Simone A. Del-Ducca Barbedo - Serviço de Informação e Documentação (SID)
REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:
Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)
Marilúcia Santos Melo Cid - Serviço de Informação e Documentação (SID)
Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)
EDITORAÇÃO ELETRÔNICA:
Viveca Sant´Ana Lemos - Serviço de Informação e Documentação (SID)

AGRADECIMENTOS

Agradeço a contribuição do Dr. Rafael Duarte Coelho dos Santos, pelas aulas de Mineração de Dados e pelas discussões metodológicas que propiciaram o desenvolvimento desse trabalho. Agradeço também ao Dr. Reinaldo Roberto Rosa pela orientação e discussões metodológicas. Por fim, também expresso minha gratidão à Dra. Débora Fernandes Calheiros por apresentar o problema e ceder gentilmente os dados analisados neste trabalho.

RESUMO

Com a modernização da tecnologia para coleta de dados, na forma de séries temporais, novos desafios metodológicos, para análise e interpretação dos padrões de variabilidade medidos, são crescentes. Em geral, desafios analíticos e interpretativos ocorrem devido aos diversos problemas de coleta de dados, quando sujeitos às variações extremas no meio ambiente, envio e recebimento de dados por satélite e problemas inerentemente instrumentais. Grande parte das séries temporais coletadas, a partir de experimentos em sistemas naturais, resulta em conjuntos de séries parciais curtas compostas por poucos pontos, comprometendo assim a qualidade das análises estatísticas convencionais. Neste contexto, este trabalho teve como principal objetivo analisar uma série temporal do ciclo de cheia do rio Paraguai de 107 anos de dados coletados diariamente, mas subdividida em séries temporais curtas de 260 pontos (época de cheia) e de 105 pontos (época de seca). Para isso usamos uma nova metodologia denominada Análise Espectral Gradiente (GSA, do inglês *Gradient Spectral Analysis*) que conjuga duas técnicas matemáticas, conhecidas como Análise de Padrões-Gradiente (GPA, do inglês *Gradient Pattern Analysis*) e Análise Multiresolução por Wavelets (WMA, do inglês *Wavelet Multiresolution Analysis*). O desempenho da GSA na classificação de séries temporais curtas (sub-séries) oriundas da série temporal do rio Paraguai foi satisfatório, indicando que a metodologia pode ser robusta em futuras aplicações de classificação e monitoramento das variáveis físicas e limnológicas medidas.

EXPLORATORY STUDY OF THE VARIABILITY PATTERN OF PARAGUAY RIVER'S FLOOD CYCLE

ABSTRACT

With the modernization of data collecting technology, in the time series, new methodological challenges for analysis and interpretation of patterns of variability measured, have been growing. In general, analytical and interpretive challenges occur due to several problems in data collection, when subject to extreme variations in the environment, sending and receiving data from satellites and problems inherently instrumentals. Most of the time series collected from experiments in natural systems, results in partial sets of series composed of a few short measures, thus compromising the quality of the statistical analyzes. In this context, this work had as main objective to analyze a series of flood cycle of the Paraguay River from 107 years of data collected daily, but divided into short series of 260 points (flood season) and 105 points (the dry season). For this we use a new methodology called Gradient Spectral Analysis (GSA), which combines two mathematical techniques known as Gradient Pattern Analysis (GPA) and Wavelets Multiresolution Analysis (WMA). The performance of the GSA in the classification of short time series (subseries) derived time series of Paraguay River was satisfactory, indicating that the method can be robust to future applications of classification and monitoring of physical variables and limnological measures.

SUMÁRIO

1 INTRODUÇÃO 17
1.1 Objetivo 20
1.2 Organização do trabalho 20

2 FERRAMENTAS MATEMÁTICO-COMPUTACIONAIS 23
2.1 *Wavelets* 23
2.2 Análise de Multirresolução por *Wavelets* (WMA) 23
2.3 Análise de Padrões-Gradiente (GPA) para ST 25
2.3.1 Coeficiente de Assimetria Gradiente 28
2.4 Análise Espectral Gradiente (GSA) 29
2.4.1 Algoritmo GSA 30

3 DESCRIÇÃO DOS DADOS ANALISADOS 33
3.1 Séries canônicas 33
3.1.1 Padrões caóticos 33
3.1.2 Padrões estocásticos 35
3.2 Padrão do nível do Rio Pantanal 39

4 METODOLOGIA 41

5 RESULTADOS 45

6 CONCLUSÃO 51

REFERÊNCIAS BIBLIOGRÁFICAS 53

LISTA DE FIGURAS

Figura 1.1 – Área do Pantanal no Brasil e na América do Sul..... 17
Figura 1.2 – Exemplo de série temporal com descontinuidades. 19
Figura 2.1 – Filtro de Quadratura Espelhado (QMF). 24
Figura 2.2 – Componentes do sinal reconstruído. 25
Figura 2.3 – Três exemplos de perfis de amplitudes compostas de 100 pontos e seus respectivos padrões-gradientes. 26
Figura 2.4 – Metodologia para mapear a ST de tamanho N numa matriz quadrada de ordem $\sqrt{N} x \sqrt{N}$. No exemplo temos uma ST de 1024 pontos distribuídos em uma matriz. 27
Figura 2.5 – Passos para calcular o coeficiente de assimetria gradiente. 29
Figura 3.1 – ST curta do Mapa Logístico: (a) 260 pontos e (b) 105 pontos. 34
Figura 3.2 – ST curta do Mapa de Hénon: (a) 260 pontos e (b) 105 pontos. 34
Figura 3.3 – ST curta do Modelo de Lorenz: (a) 260 pontos e (b) 105 pontos. 35
Figura 3.4 – Exemplo de ST curta estocástica: $P(\omega) \sim \omega^{-3/2}$. 37
Figura 3.5 – ST curta com padrão fBm $\omega^{-3/2}$: (a) 260 pontos e (b) 105 pontos. ... 38
Figura 3.6 – ST curta com padrão fBm $\omega^{-5/3}$: (a) 260 pontos e (b) 105 pontos. 38
Figura 3.7 – ST curta com padrão fBm ω^{-2}: (a) 260 pontos e (b) 105 pontos. 39
Figura 3.8 – Perfil de medida diária do nível do Rio Paraguai em Ladário-MS entre 1900 e 2008. Medidas obtidas pela Marinha do Brasil. 40
Figura 4.1 – Exemplo de uma ST anual do nível do Rio Paraguai. 41
Figura 4.2 – Ciclos de cheia e seca considerados neste trabalho..... 42
Figura 4.3 – Exemplo de ST de cheia anual normalizada. 43
Figura 4.4 – Exemplo de ST de seca anual normalizada. 43
Figura 5.1 – Exemplo de resultado da GSA: flutuação média do espectro gradiente 45
Figura 5.2 – ST completa (107 anos) de nível do rio Paraguai (parte superior), ... 47
juntamente com os tipos de variabilidade para cada época (parte inferior). 47
Figura 5.3 – ST parcial (3 anos) de nível do rio Paraguai (parte superior), juntamente com os tipos de variabilidade para cada época (inferior).48
Figura 5.4 – Distribuição de padrões caóticos e estocásticos nas épocas de (a) Seca e (b) Cheia..... 48

LISTA DE TABELAS

Tabela 5.1 – padrões de variabilidade (1: Logístico; 2: Hénon; 3: Lorenz; 4: Estocástico 46

LISTA DE SIGLAS E ABREVIATURAS

INPE	Instituto Nacional de Pesquisas Espaciais
ST	Série(s) temporal(is)
FFT	*Fast Fourier Transform* (Transformada Rápida de Fourier)
WMA	*Wavelet Multiresolution Analysis* (Análise de Multirresolução por *Wavelets*)
GPA	*Gradient Pattern Analysis* (Análise de Padrões-Gradiente)
GSA	*Gradient-Spectra Analysis* (Análise Espectral dos Gradientes)
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
LAC	Laboratório Associado de Computação e Matemática Aplicada
MS	Mato Grosso do Sul
CWT	*Continuous Wavelet Transform* (Transformada Contínua de Wavelet)
ICWT	*Inverse Continuous Wavelet Transform* (Transformada Inversa Contínua de Wavelet)
DWT	*Discrete Wavelet Transform* (Transformada Discreta de Wavelet)
IDWT	*Inverse Discrete Wavelet Transform* (Transformada Inversa Discreta de Wavelet)
QMF	*Quadrature Mirror Filters* (Filtros de Quadratura Espelhados)

1 INTRODUÇÃO

O Pantanal é uma savana tropical inundada localizada na América do Sul ocupando uma área de cerca de 137.000 km^2 (aproximadamente 450 por 300 km) em território brasileiro (HAMILTON et al., 1996). A maior parte da região está situada no Brasil, com pequenas áreas (cerca de 20% do total) na Bolívia e Paraguai (Figura 1.1).

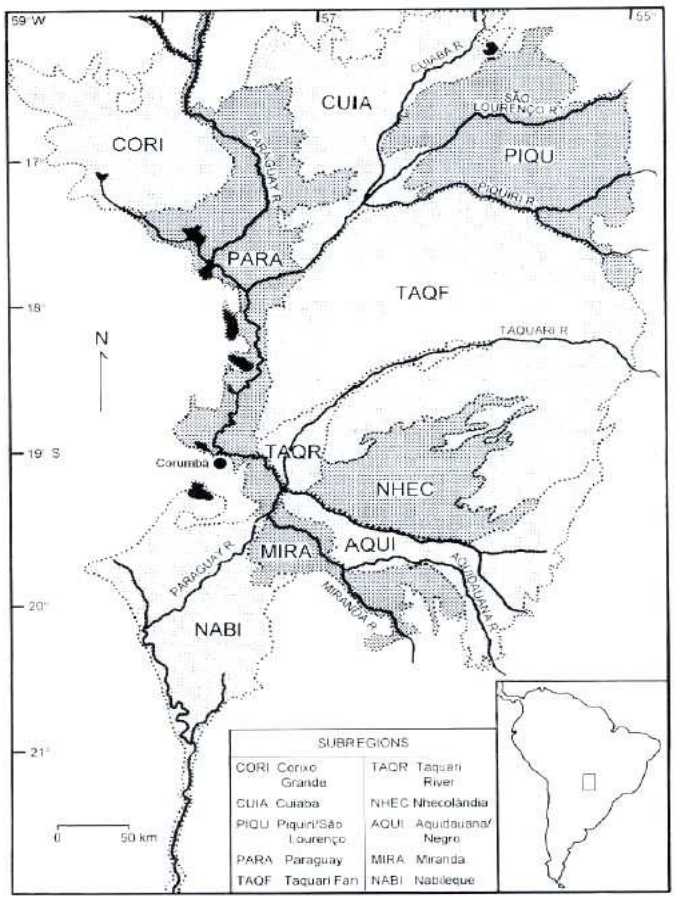


Figura 1.1 – Área do Pantanal no Brasil e na América do Sul.
Fonte: HAMILTON et al. (1996).

A região do Pantanal é caracterizada como um típico sistema com dinâmica de inundações. Um dos seus principais rios, o rio Paraguai, é de fundamental influência nessa dinâmica de inundações. O estudo da hidrologia, da biogeoquímica e da ecologia de regiões tropicais sujeitas a inundações perpassa pelo entendimento dos padrões de variabilidade entre áreas secas e inundadas sazonalmente (HAMILTON et al., 1996; RICHEY et al., 1991; BARTLETT e HARRISS, 1993; e JUNK, 1993). Além disso, tal entendimento também é fundamental para previsão de viabilidade e impactos ambientais de projetos no sentido de regular essas inundações para um determinado fim (HAMILTON et al., 1996; BUCHER et al., 1993).

O clima da região é tropical com uma estação chuvosa bem característica (BRASIL, 1979). A precipitação pluviométrica anual está entre 1.000 a 1.700 mm, com chuvas mais concentradas entre novembro e março.

Esse trabalho analisa os dados de nível diário do rio Paraguai que foram coletados em Ladário-MS e cedidos originalmente pelo VI Distrito Naval – Serviço de Sinalização Náutica do Oeste, Marinha do Brasil. Esses dados formam uma série temporal. Neste trabalho usamos a definição de série temporal segundo DANTAS (2008), a qual a partir deste ponto será denominada de ST, tanto no singular, quanto no plural.

Conforme DANTAS (2008), dois dos principais problemas em ST reais são o tamanho da série e a não-estacionariedade dos processos geradores da mesma. O primeiro ocorre em decorrência de descontinuidades significativas (ver Figura 1.2) causadas principalmente por falhas diversas nos equipamentos de coleta de dados. Já a segunda, decorre da variabilidade complexa comum às ST oriundas de dados coletados de variáveis físicas ambientais (meteorológicas, hidrológicas, oceanográficas ou limnológicas). Essa variabilidade é interpretada por BARBOSA

(2006) como “regime de variação que compreende antipersistência, persistência e intermitência no domínio temporal”.

Figura 1.2 – Exemplo de série temporal com descontinuidades.
Fonte: Adaptado de INPE (2008a).

Na Figura 1.2 são apresentados segmentos de uma série de medidas diárias da temperatura da água de um reservatório de Furnas (INPE, 2008(a)). Note que, devido a problemas instrumentais ou de transmissão de dados, o conjunto de medidas destacado em vermelho possui apenas 75 medidas de um conjunto total de 384 (mesmo numa resolução horária teríamos 1800 medidas na região de destaque de um total de 9216).

Essas características demandam um tratamento apropriado não convencional para análise da ST, já que a ST precisa ter um tamanho mínimo para um tratamento convencional robusto e a maioria das aplicações convencionais não lida corretamente com a não-estacionariedade de uma série, segundo MORETIN (1999).

Assim, a característica inerente à ST curta de ter o tamanho da ordem de até 10^3 pontos de medida descredencia a aplicação de métodos convencionais, que exigem uma ST com número de pontos $N >> 10^3$.

1.1 Objetivo

Analisar através da metodologia GSA padrões de variabilidade de ST curtas obtidas dos ciclos de cheia e seca do Rio Pantanal comparando com modelos estocásticos e determinísticos, classificando esses padrões, com diferentes graus de autocorrelação, auto-afinidade, intermitência e persistência.

1.2 Organização do trabalho

Os demais capítulos deste trabalho estão organizados da seguinte maneira:

- Capítulo 2: traça um breve panorama dos aspectos gerais das principais ferramentas matemáticas utilizadas: as *Wavelets*[1], a Análise de Multirresolução, a Análise de Padrões-Gradiente para ST e a Análise Espectral Gradiente.
- Capítulo 3: descreve os dados analisados neste trabalho.
- Capítulo 4: aborda a metodologia da análise.
- Capítulo 5: apresenta a aplicação das ferramentas nos dados e discute os principais resultados deste trabalho.

[1] Uma vez que há controvérsias para o termo em português (ondeletas x ondaletas), convencionamos utilizar o termo em inglês.

- Capítulo 6: organiza as conclusões provenientes deste projeto, enfatizando a contribuição científica das mesmas. Finalmente, são sugeridas perspectivas de pesquisas futuras.

2 FERRAMENTAS MATEMÁTICO-COMPUTACIONAIS

Neste capítulo descrevemos brevemente as técnicas matemáticas utilizadas neste trabalho, tais como: as *Wavelets*, a Análise de Multirresolução por *Wavelets* (WMA), a Análise de Padrões-Gradiente (GPA) para ST e a Análise Espectral Gradiente (GSA).

2.1 *Wavelets*

As *wavelets* são "alternativas a outros sistemas de funções usados como bases para representação de funções pertencentes a certos espaços, como os senos e cossenos, polinômios ortogonais, funções Walsh, etc.", segundo MORETTIN (1999). Elas constituem atualmente uma das ferramentas mais discutidas, avaliadas e utilizadas no processamento digital de sinais.

A Transformada de *Wavelet* constitui uma ferramenta moderna que permite a unificação de várias técnicas de análise e processamento. É uma ferramenta que decompõe dados, funções ou operadores em diferentes componentes de freqüência (ou escala) e então estuda cada componente com uma resolução coerente com sua escala. Para maiores detalhes históricos e de aplicações, consulte DANTAS (2008).

2.2 Análise de Multirresolução por *Wavelets* (WMA)

A Análise de Multirresolução pode ser efetuada através da decomposição e da reconstrução multi-nível de uma ST. Estas são processos matemáticos que correspondem à Transformada Discreta de *Wavelet* e à Transformada Inversa Discreta de *Wavelet*, respectivamente.

Através de filtros de decomposição passa-baixa (L) e passa-alta (H), juntamente com seus respectivos filtros de reconstrução (L'e H'), a WMA forma um sistema que é chamado Filtro de Quadratura Espelhado[2] (QMF, do inglês *Quadrature Mirror Filters*). Veja esquema do QMF na Figura 2.1.

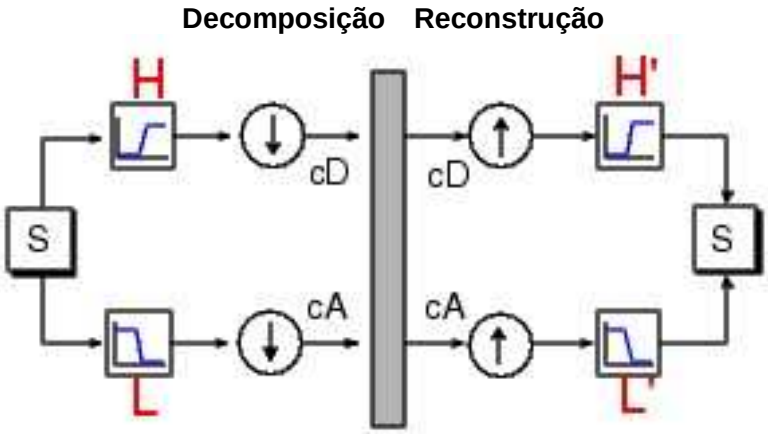


Figura 2.1 – Filtro de Quadratura Espelhado (QMF).

Esse sistema, segundo DANTAS (2008), possibilita a obtenção de uma aproximação real do sinal original (baixas freqüências), que tem o mesmo tamanho do sinal original S e, através de um processo análogo, podemos fazer a reconstrução do primeiro nível de detalhes (altas freqüências). Os componentes de detalhes e os de aproximações reconstruídos são os verdadeiros constituintes do sinal original. De fato, $A_1 + D_1 = S$. Podemos estender esse raciocínio para uma análise multinível. Veja Figura 2.2.

[2] O filtro de quadratura espelhado é um filtro muito utilizado para reconstruir um sinal a partir de seus coeficientes *wavelets* (MALLAT, 1989b).

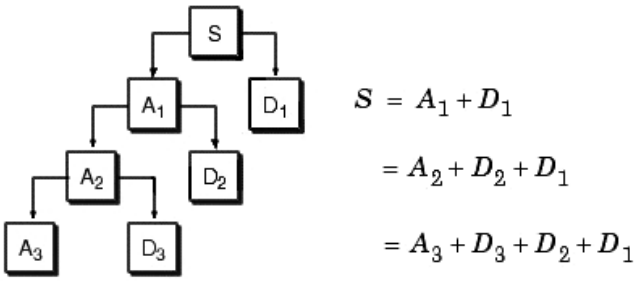


Figura 2.2 – Componentes do sinal reconstruído.

Neste trabalho estamos interessados em classificar assimetrias em ST, sendo algumas não-estacionárias. Portanto, é conveniente fazer a análise multi-nível em componentes de aproximação, pois nessas componentes a informação de persistência se mantém.

2.3 Análise de Padrões-Gradiente (GPA) para ST

O principal objetivo da GPA é quantificar assimetrias em escalas locais e globais de um dado perfil temporal, espacial ou espaço-temporal, por meio de uma operação computacional que caracteriza padrões – através das medidas de pequenas e grandes amplitudes em tais padrões – como grades gradientes (ou uma seqüência de matrizes) (ROSA et al., 1999; ROSA et al., 2000; ASSIREU et al., 2002; ROSA et al., 2003; BARONI et al., 2006b; ROSA et al., 2007). Essa grade gradiente é representada por uma matriz denominada, matriz das amplitudes:

$$M_A = M_A(1,1),...,M_A(i,j),...,M_A(\sqrt{N},\sqrt{N}) \mid i,j \in I \quad e \quad M_A \in \Re, \tag{2.1}$$

onde N é o tamanho da série temporal. A matriz quadrada M_A, possui dimensões espaciais (x,y), discretizadas em $\sqrt{N}x\sqrt{N}$ pontos, com $1 \geq i \geq \sqrt{N}$ e $1 \geq j \geq \sqrt{N}$.

A flutuação espacial do padrão global $M_A(i, j)$, para um dado instante t, pode ser caracterizada através do campo vetorial gradiente $G_t = \nabla[M_A(x, y)]_t$. Uma flutuação espacial local, entre um par instantâneo de intensidades e pertencentes ao padrão global, é caracterizada por seu vetor gradiente, definido entre cada par de pontos da grade bi-dimensional. Nesta representação, os valores relativos entre as amplitudes (que determinam a norma e a orientação de cada vetor) são mais relevantes do que os seus valores absolutos.

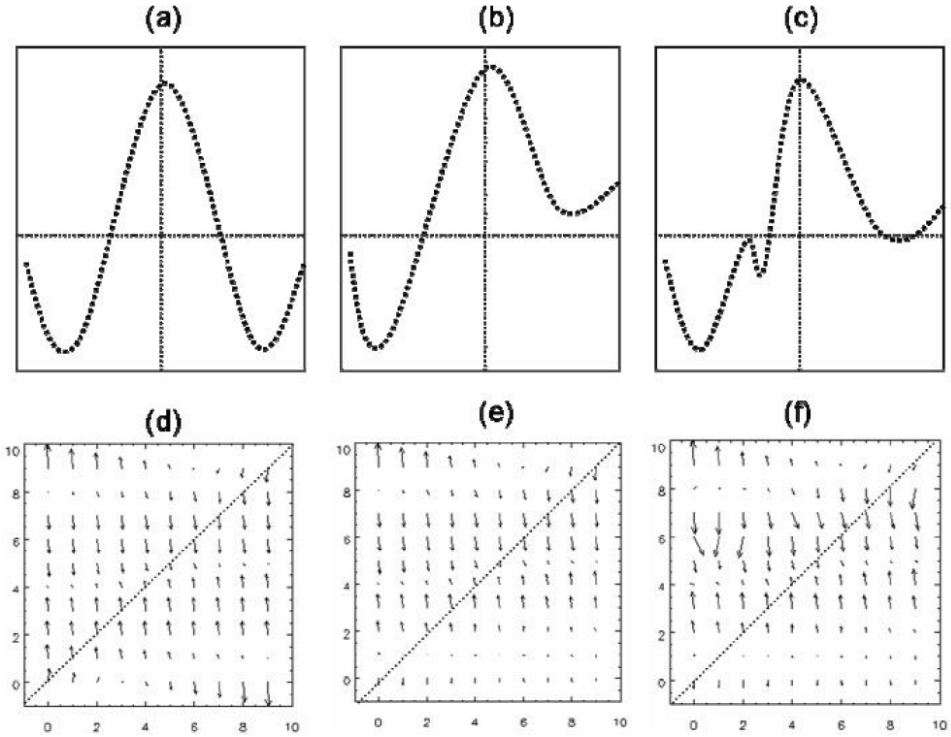


Figura 2.3 – Três exemplos de perfis de amplitudes compostas de 100 pontos e seus respectivos padrões-gradientes.

Nesse trabalho usamos o conceito de simetria para análise em ST, conforme ASSIREU et al,. (2002). O conceito de simetria gradiente é ilustrado na Figura 2.3. A Figura 2.3(a) mostra um perfil totalmente simétrico em relação ao eixo vertical.

O perfil simétrico da Figura 2.3(a) é composto por 100 pontos de modo que a sua matriz quadrada tem a forma 10 x 10 (veja a Figura 2.4, onde é mostrado como é feito o mapeamento dos valores da ST para o formato de matriz). O padrão gradiente correspondente a esse perfil é mostrado na Figura 2.3(d). Levando em consideração um eixo diagonal nesse campo gradiente, podemos observar que para cada vetor local $\underline{v}$ da grade vai existir um vetor correspondente $-\underline{v}$ com o mesmo módulo, mas com fase oposta – denominados vetores simétricos. Assim, se removermos os pares simétricos para quantificar assimetria, ao final da operação, não haverá vetores remanescentes no padrão gradiente. As Figuras 2.3(b), 2.3(c) são exemplos de perfis assimétricos e seus padrões-gradientes podem ser visualizados nas Figuras 2.3(e), 2.3(f), respectivamente.

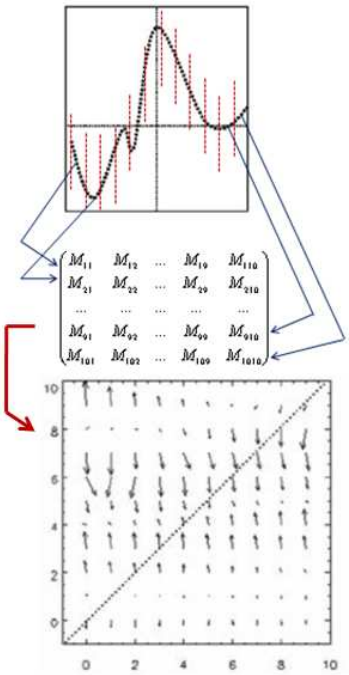


Figura 2.4 – Metodologia para mapear a ST de tamanho N numa matriz quadrada de ordem $\sqrt{N} x \sqrt{N}$. No exemplo temos uma ST de 1024 pontos distribuídos em uma matriz.

A Figura 2.4 demonstra como é feito o mapeamento da ST para a matriz correspondente e o seu respectivo padrão-gradiente. Como exemplo, uma ST contendo 100 pontos corresponde a uma matriz de tamanho $10x10$, onde cada linha da matriz é um grupo de 10 pontos seqüenciais da ST, tomados da esquerda para a direita. Como mostrado em (ASSIREU et al., 2002), os valores de G_A calculados a partir dessas matrizes não dependem da direção da ST como são tomados (da direita para a esquerda ou vice-versa).

2.3.1 Coeficiente de Assimetria Gradiente

A medida do gradiente global de vetores assimétricos, denominada aqui de *Coeficiente de Assimetria Gradiente*[3] G_A, é dado por:

$$G_A = \frac{N_C - N_V}{N_V}, \tag{2.2}$$

onde $N_C > N_V$. N_V é o número total de vetores assimétricos remanescentes após a remoção dos pares simétricos e N_C é o número de conexões entre esses vetores (gradientes locais). Como o perfil mostrado na Figura 2.3(a) é totalmente simétrico, o gradiente assimétrico não existe e seu coeficiente de assimetria gradiente é nulo ($G_A \equiv 0$). Esse operador computacional mede a quebra de simetria de uma dada grade de flutuação e tem sido usado em várias aplicações (ROSA et al., 1999; ROSA et al., 2000; ASSIREU et al., 2002; ROSA et al., 2003; BARONI et al., 2006a,b; ROSA et al., 2007). A Figura 2.5 mostra um diagrama que ilustra a seqüência de passos para o cálculo do coeficiente de assimetria para uma ST.

[3] Esse é um dos quatro momentos gradientes apresentados por (ROSA et al., 2003). É importante ressaltar que esses momentos possuem a propriedade de serem invariantes globalmente em relação à rotação da grade e modulação da amplitude.

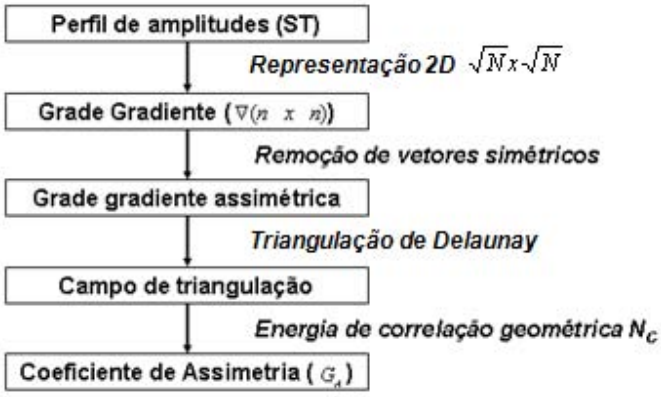


Figura 2.5 – Passos para calcular o coeficiente de assimetria gradiente.

Cada gradiente local é obtido pela computação de

$$\frac{d(M_A(i,j))}{di} \text{ e } \frac{d(M_A(i,j))}{dj} \tag{2.3}$$

da matriz, que representam as diferenças de amplitudes do perfil nas direções i (linha) e j (coluna), respectivamente. O espaçamento entre os pontos em cada direção é assumido como um. Uma rotina protótipo desse tipo de gradiente e a rotina clássica do campo de triangulação de Delaunay podem ser encontradas em (ASSIREU et al., 2002) e (BARBER, 1996), respectivamente. Para maiores detalhes, DANTAS (2008).

2.4 Análise Espectral Gradiente (GSA)

A GSA consiste em uma seqüência ordenada de quatro operações (implementadas computacionalmente) sobre uma dada ST: (i) determinação da Escala Máxima de Coerência da ST; (ii) representação multirresolução da ST; (iii) cálculo da Potência do Coeficiente de Assimetria Gradiente e (iv) obtenção do Espectro-

Gradiente e da Medida de Flutuação do Espectro-Gradiente. Uma explicação detalhada de cada uma dessas operações está contida em DANTAS (2008).

2.4.1 Algoritmo GSA

O algoritmo da GSA é constituído pelas quatro operações computacionais explicadas acima. Abaixo, demonstramos essas operações em forma de esquema.

Primeira operação: escolher a λ_{mc} da ST.

1. Ler a ST de um arquivo;
2. Calcular o diagrama tempo-escala da ST;
3. Calcular a variância de cada escala;
4. Retornar a escala de maior variância.

Segunda operação: fazer a Análise de Multirresolução da ST.

1. Decompor a ST discretamente (multi-nível por aproximação);
2. Reconstruir a ST discretamente (multi-nível por aproximação).

Terceira operação: calcular $< G_A >$ e G_{POT}.

1. Calcular o G_A janelado por λ_{mc} da ST em cada resolução;
2. Calcular a G_{POT} para cada resolução;
3. Retornar os valores de G_{POT}.

Quarta operação: gerar o Espectro-Gradiente e sua f_{eg}.

1. Imprimir na tela em forma de gráfico os valores retornados do script de G_{POT};

2. Calcular flutuação: comparar a medida de G_{POT} com o $< G_{POT} >$ do próprio sinal;

3. Salvar os valores das flutuações, do desvio padrão e da variância.

No próximo capítulo descrevemos os dados analisados neste trabalho.

3 DESCRIÇÃO DOS DADOS ANALISADOS

Neste capítulo descrevemos os dados selecionados em função das análises previstas no projeto. Nosso estudo de ST curtas consiste em classificar ST curtas dos ciclos de cheia e seca do Rio Paraguai.

Todas as séries canônicas apresentadas neste trabalho foram obtidas por meio de simulação computacional obedecendo ao respectivo modelo do padrão escolhido. Já as séries ambientais foram obtidas através do grupo de pesquisa parceiro, Embrapa-MS. Com o intuito de nivelar o domínio dos dados, todas as séries selecionadas nos períodos de cheia e seca são constituídas por 260 e 105 medidas de amplitudes, respectivamente, normalizadas entre zero (0) e um (1). Portanto, com o intuito de obtermos a classificação, para cada tipo canônico temos exemplos de ST com 260 e 105 medidas.

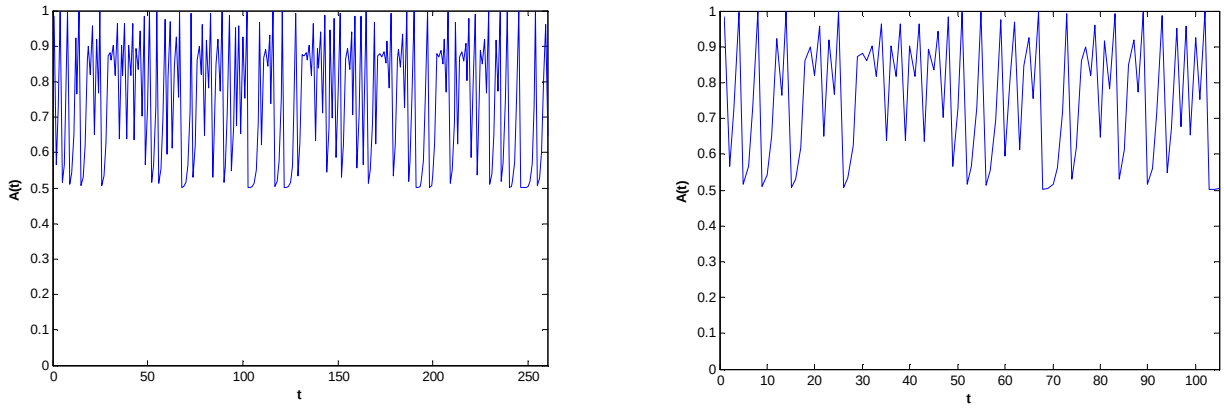
3.1 Séries canônicas

3.1.1 Padrões caóticos

Nesta pesquisa escolhemos três modelos para representar os padrões caóticos: o Mapa Logístico (unidimensional), o Mapa de Hénon (bidimensional) e o Modelo de Lorenz (tridimensional). Para maiores detalhes sobre esses modelos, veja DANTAS (2008).

Mapa Logístico

Na Figura 3.1 é possível visualizar um exemplo do padrão do Mapa Logístico, com $\rho = 4{,}0$ e a condição inicial para X_0 sendo um valor aleatório entre 0 e 1, com visualização de (a) 260 e (b) 105 pontos.

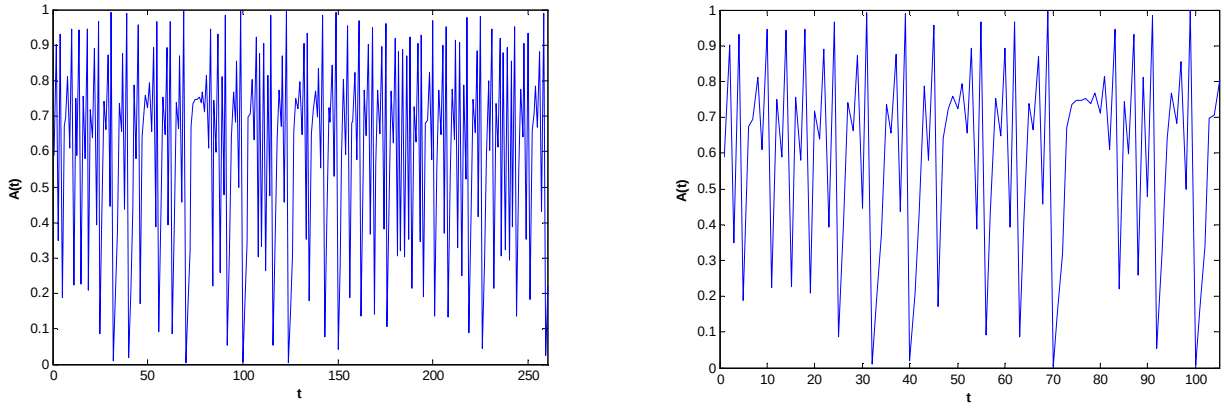


(a) (b)

Figura 3.1 – ST curta do Mapa Logístico: (a) 260 pontos e (b) 105 pontos.

Mapa de Hénon

Na Figura 3.2 é possível visualizar um exemplo do padrão do Mapa de Hénon, com os parâmetros canônicos $a = 1{,}4$ e $b = 0{,}3$ típicos de um processo com comportamento caótico e as condições iniciais para X_0 e Y_0 sendo valores aleatórios entre 0 e 1, com visualização de (a) 260 e (b) 105 pontos.

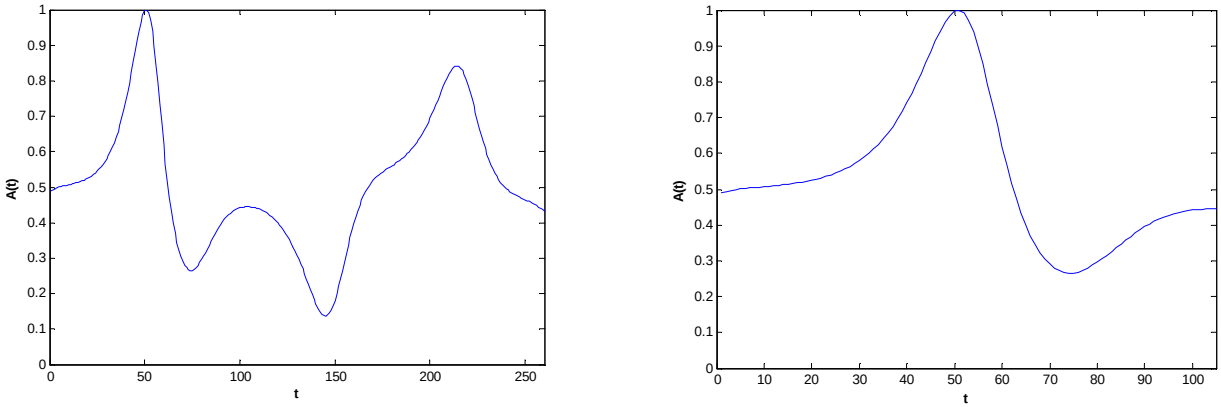


(a) (b)

Figura 3.2 – ST curta do Mapa de Hénon: (a) 260 pontos e (b) 105 pontos.

Modelo de Lorenz

Em nossa pesquisa utilizamos os seguintes valores usuais (PEITGEN et al., 1992): $\sigma = 10$, $\beta = 28$ e $\gamma = 8/3$. Para as condições iniciais de X_0, Y_0 e Z_0, utilizamos valores aleatórios entre 0 e 1 (ver exemplo de uma ST gerada, na Figura 3.3, com visualização de (a) 260 pontos e (b) 105 pontos).



(a) (b)

Figura 3.3 – ST curta do Modelo de Lorenz: (a) 260 pontos e (b) 105 pontos.

Portanto, com base nos três modelos caóticos apresentados nesta seção, construímos algoritmos computacionais (em MATLAB®) com os quais geramos exemplos de ST com padrões de variabilidade caótica.

3.1.2 Padrões estocásticos

Os processos estocásticos estão associados ao comportamento de variáveis aleatórias cuja dinâmica é regida por uma grande quantidade de graus de liberdade e, portanto, são chamados de processos de alta dimensão (DANTAS, 2008).

O espectro de um processo estocástico pode ser representado por uma série discreta de freqüências e a sua função aleatória correspondente pode ser escrita como uma série discreta para tempos $t_i = i\Delta t$ com $i = 1,...,M$, permitindo que a ST

possa ser escrita como uma simples superposição de oscilações harmônicas, dada por (OSBORNE e PROVENZALE, 1989):

$$A(t_i) = \sum_{k=1}^{M/2} [P(\omega_k)\Delta\omega]^{1/2} \cos(\omega_k t_i + \phi_k), \quad i = 1,...,M \text{ ,} \tag{3.1}$$

onde $\omega_k = k\Delta\omega$, com $k = 1,...,M/2$, $\Delta\omega = 2\pi / Mt$, ϕ_k sendo as fases escolhidas aleatoriamente e M sendo o número de pontos da ST.

A variabilidade em ST geradas por tais padrões complexos é do tipo persistente, como pode ser visto na Figura 3.4. Esse padrão foi gerado através da Equação 3.1, utilizando o algoritmo de MALAMUD e TURCOTE (1999) com $\alpha = \frac{3}{2}$ (uma Lei de Potência típica de padrões estocásticos de alta dimensão). Note que, na Equação 3.1, $P(\omega) = 1/\omega^{\alpha}$.

Conforme DANTAS (2008), o conceito de Lei de Potência, associado a um espectro de energias, permite caracterizar um padrão de variabilidade do tipo turbulento, isto é, cujas densidades espectrais são proporcionais a $\frac{1}{\omega^{\alpha}}$ com α assumindo diferentes valores a partir do tipo de turbulência considerado, num intervalo de algumas ordens de grandeza da freqüência. Sinais gerados por processos estocásticos do tipo $\frac{1}{\omega^{\alpha}}$ são encontrados na física, meteorologia, biologia, engenharia, economia, etc. Para maiores detalhes, consultar DANTAS (2008).

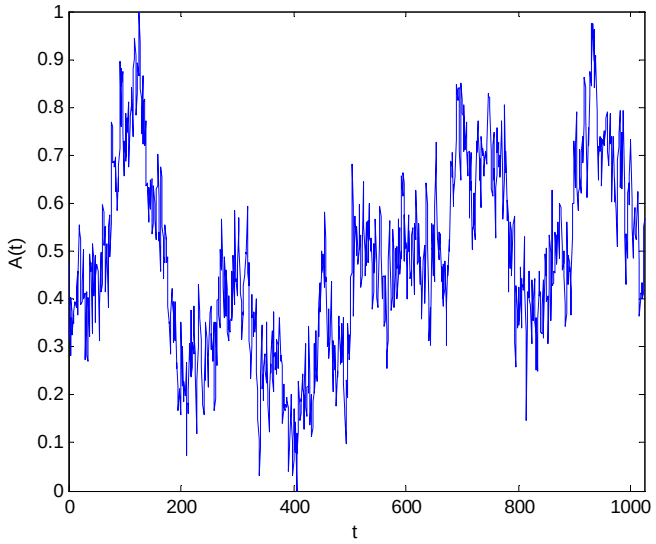


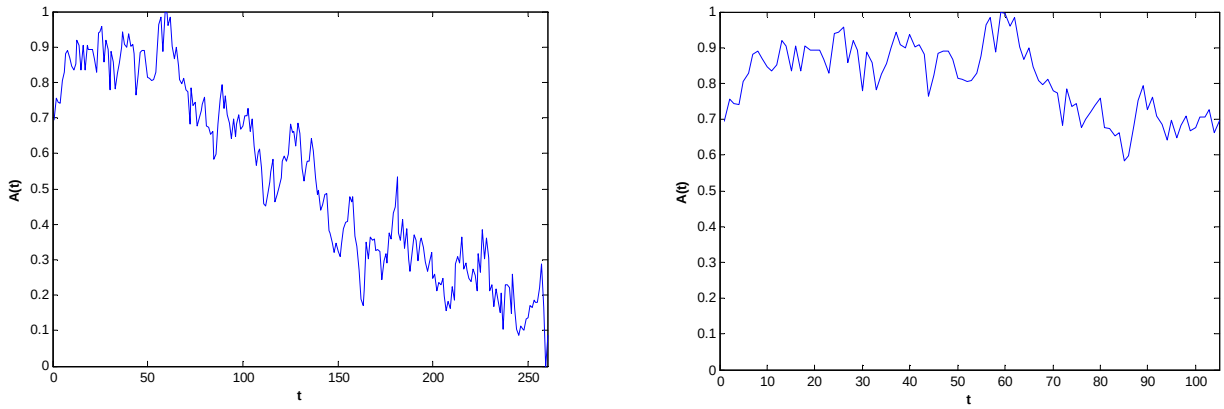
Figura 3.4 – Exemplo de ST curta estocástica: $P(\omega) \sim \omega^{-3/2}$.

A equação 3.4 permite a simulação de um processo físico, muito comum, chamado movimento fracionário Browniano (fBm). O fBm foi introduzido por MANDELBROT e VAN NESS (1968) como um meio de representar processos estocásticos não estacionários que exibem dependência de longo alcance e/ou que têm propriedade de auto-similaridade.

Dois aspectos importantes devem ser levados em conta na análise do fBm: a sua não-estacionariedade, que necessita de alguma análise dependente do tempo, e a sua auto-similaridade, que necessita de alguma análise dependente da escala (FLANDRIN, 1992 e FLANDRIN, 1994). Logo, um estudo consistente do padrão fBm deve considerar técnicas de decomposição em tempo-escala. Nesse contexto, a análise por wavelets é particularmente adequada para analisar o fBm e caracterizar sua riqueza de escalas (FLANDRIN, 1992; TEWFICK e KIM, 1992 e FLANDRIN, 1994).

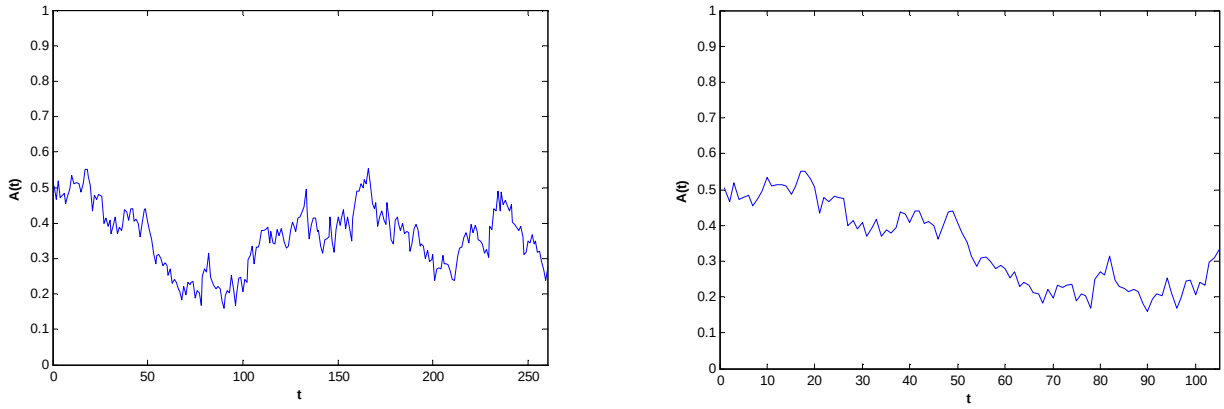
Neste trabalho, selecionamos três tipos de padrões para representar a variabilidade fBm: $\alpha = -\frac{3}{2}$, $\alpha = -\frac{5}{3}$ e $\alpha = -2$ (Figuras 3.5, 3.6 e 3.7, respectivamente).

Com o respectivo algoritmo da Equação 3.1 foram geradas séries para cada Lei de Potência apresentada. Dessa forma, juntamente com os dados caóticos descritos anteriormente, essas séries constituem o conjunto de dados "canônicos" para estudo das metodologias a serem aplicadas nas ST reais medidas no Pantanal.



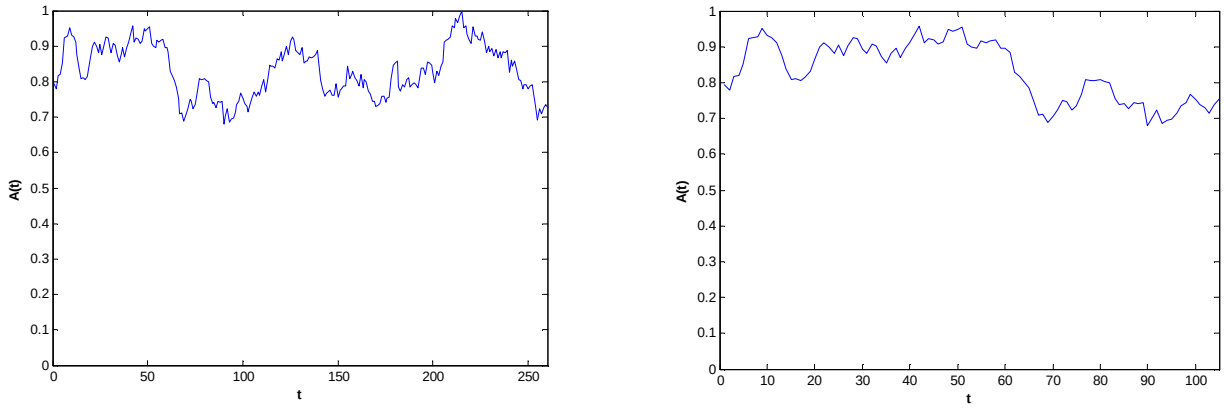
(a) (b)

12Figura 3.5 – ST curta com padrão fBm $\omega^{-3/2}$: (a) 260 pontos e (b) 105 pontos.



(a) (b)

Figura 3.6 – ST curta com padrão fBm $\omega^{-5/3}$: (a) 260 pontos e (b) 105 pontos.



(a) (b)

Figura 3.7 – ST curta com padrão fBm ω^{-2} : (a) 260 pontos e (b) 105 pontos.

3.2 Padrão do nível do Rio Pantanal

A régua centenária da Marinha do Brasil situada em Ladário-MS é uma das principais referências do ciclo de cheia ou seca no Pantanal, pois esse ciclo é fortemente influenciado pelo rio Paraguai, onde a régua está instalada. Assim, essa ST, que tem resolução diária, tem sido objeto de estudo por muitos pesquisadores no intuito de estudar a dinâmica de inundação do Pantanal. A Figura 3.8 mostra o perfil dessa ST.

O perfil mostrado na Figura 3.8 foi dividido em períodos de 365 pontos, onde cada período foi subdividido em época de cheia (260 pontos) e época de seca (105 pontos). ST curtas assim inviabilizam o uso de técnicas estatísticas para a classificação dos padrões de variabilidade observados. Tentando contornar esta limitação metodológica, aplicamos a GSA para essa análise com o objetivo de comparar os padrões de variabilidade do nível do rio Paraguai com os modelos canônicos previamente escolhidos. O GSA já havia demonstrado robustez para ST tão curtas (DANTAS, 2008).

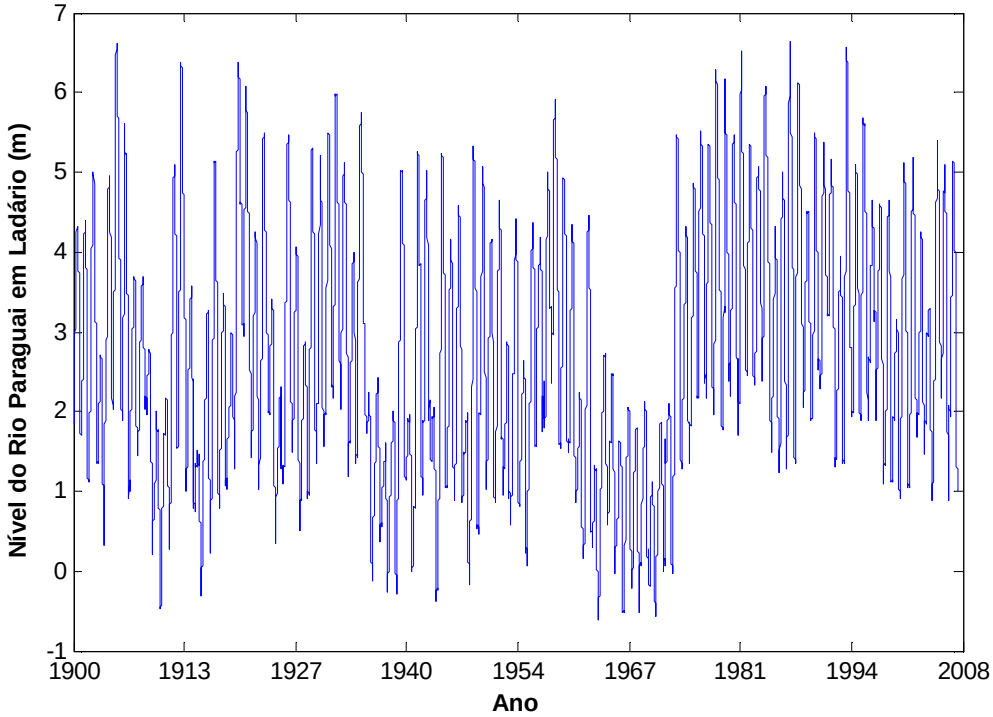


Figura 3.8 – Perfil de medida diária do nível do Rio Paraguai em Ladário-MS entre 1900 e 2008. Medidas obtidas pela Marinha do Brasil.

No capítulo a seguir, apresentamos o desenvolvimento deste trabalho relacionando os processos canônicos e ambientais descritos neste capítulo.

4 METODOLOGIA

A ST representada na Figura 3.8 é composta por medidas diárias do nível do Rio Paraguai desde o ano 1900 até o ano 2008. A fim de uniformizar o tamanho das séries, retiramos a medida adicional em anos bissextos. Ou seja, todas as séries anuais foram consideradas com 365 pontos. Veja Figura 4.1 para um exemplo de série anual. É importante ressaltar que através desta figura podemos observar que a maior parte do ano corresponde à época de cheia do que de seca.

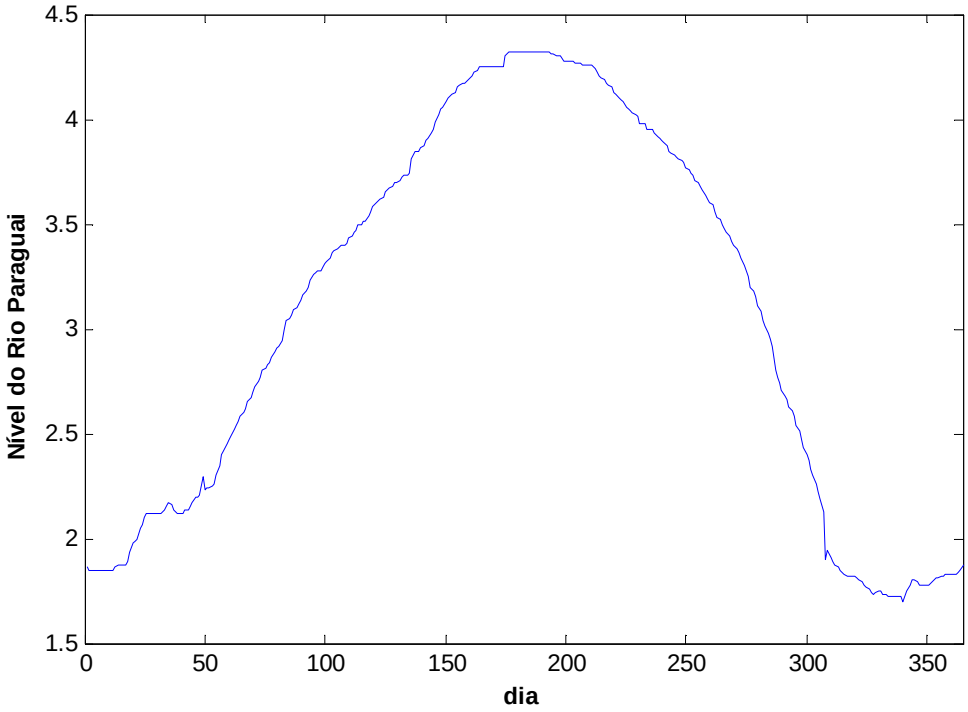


Figura 4.1 – Exemplo de uma ST anual do nível do Rio Paraguai.

Dentre as 108 ST anuais, calculamos uma média de dias do ano correspondente à fase de cheia e outra média de dias do ano correspondente à fase de seca, obten-

do como valores respectivos, 260 e 105 dias. Baseamos nossos cálculos nos momentos de inicialização e finalização da cheia.

Consideramos também a continuidade das épocas de cheia e seca. Como pode ser observado na Figura 4.1, o ciclo de cheia contínuo está contido no ano vigente, já o ciclo de seca corresponde a dois pedaços da ST: o primeiro no início do ano e o segundo, no fim. A simples junção desses dois períodos poderia, no nosso entender, embutir uma transição entre padrões de variabilidade diferente. Por conta disso, eliminamos a seca do início do ano de 1900, a seca de fim de ano do ano 2008 e a cheia de 2008, ou seja, consideramos apenas 107 ST.

Portanto, com o intuito de evitar esse viés em nossa análise, adotamos o critério de escolher o período de seca entre dois anos consecutivos, como mostra a Figura 4.2.

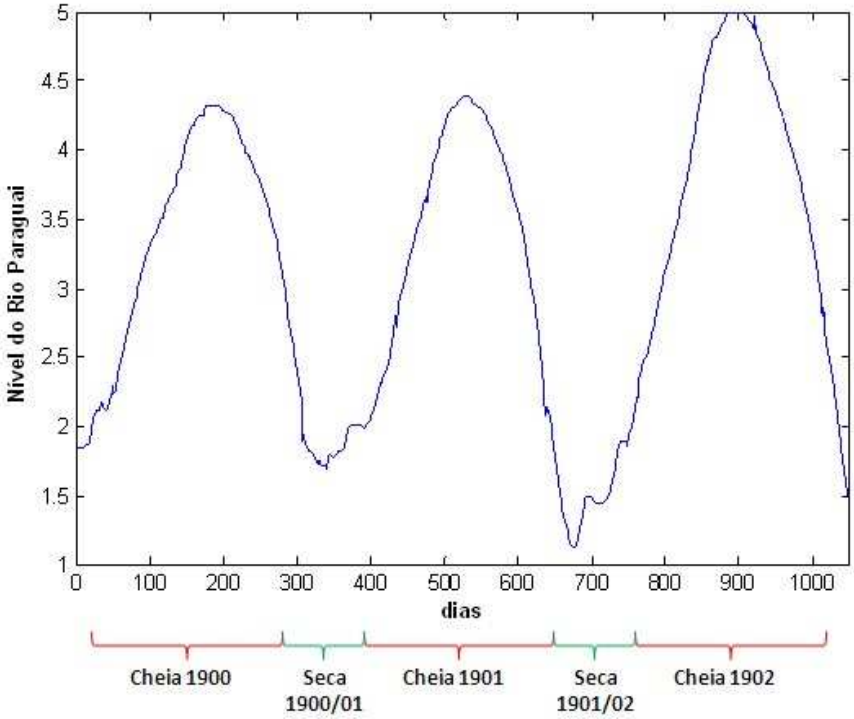


Figura 4.2 – Ciclos de cheia e seca considerados neste trabalho.

Em seguida, normalizamos a ST completa entre 0 e 1 e separamos as respectivas ST de cheia e seca (Veja as Figuras 4.3 e 4.4, respectivamente). Para cada uma delas, analisamos através do GSA sua classificação em relação às ST canônicas apresentadas no capítulo anterior.

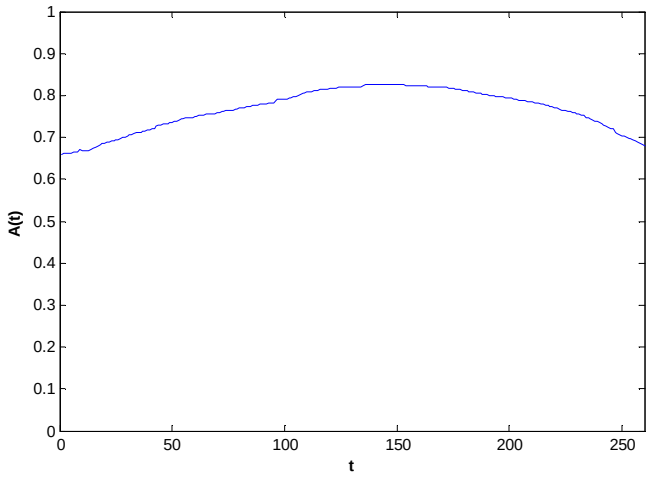


Figura 4.3 – Exemplo de ST de cheia anual normalizada.

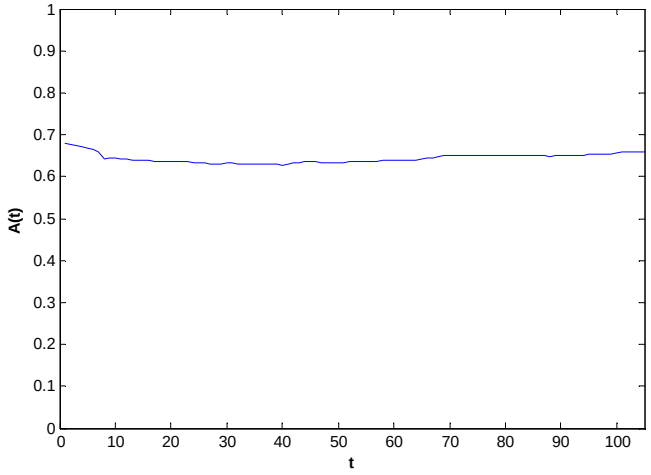


Figura 4.4 – Exemplo de ST de seca anual normalizada.

5 RESULTADOS

Este capítulo aborda os resultados obtidos com a aplicação da GSA descrita no capítulo anterior nas ST canônicas e ambientais para a respectiva classificação.

Para usar a GSA, tivemos que utilizar a *wavelet* biortogonal com ordem de reconstrução igual a dois e ordem de decomposição igual a quatro (bior2.4), devido ao tamanho da ST. Para essas ST com esses tamanhos, essa *wavelet* gerou apenas dois níveis de resolução *wavelet*. Para maiores detalhes sobre a escolha dessa *wavelet* e seus parâmetros, consultar DANTAS (2008).

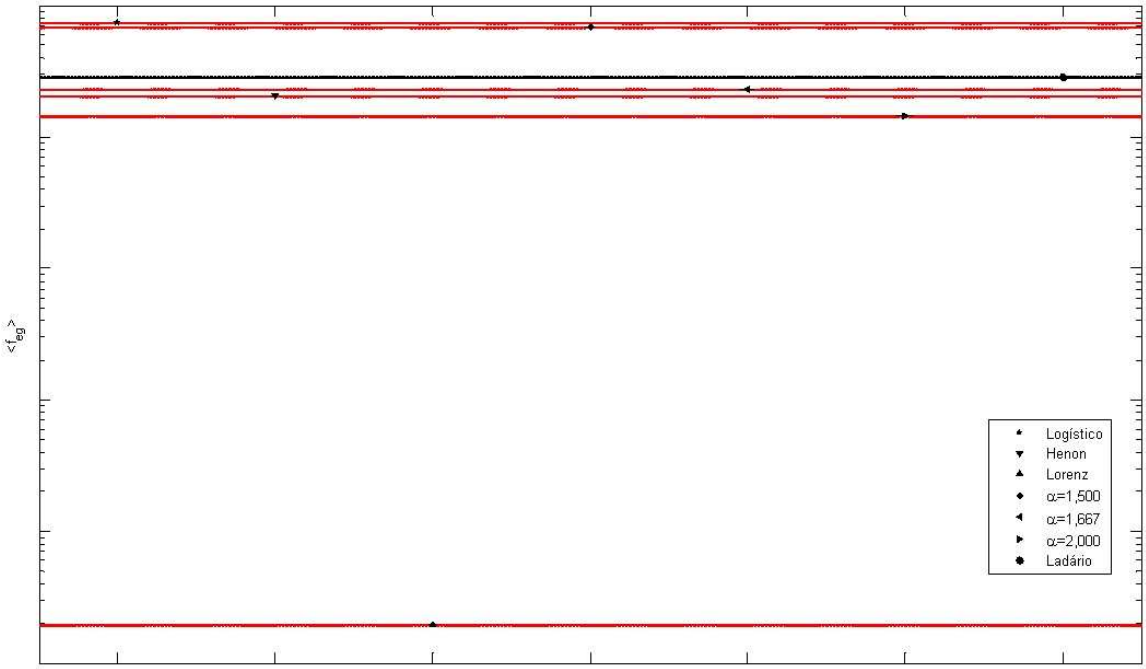


Figura 5.1 – Exemplo de resultado da GSA: flutuação média do espectro gradiente em uma época de seca.

Na Figura 5.1, podemos visualizar a classificação da GSA de uma das ST de seca. Neste exemplo, a classificação aponta para o padrão estocástico com $\alpha = -\frac{5}{3}$.

Ou seja, a proximidade do resultado da GSA, em relação aos demais, entre os padrões de variabilidade da ST e o padrão estocástico, nos permite inferir que o comportamento neste período se assemelha a esse comportamento estocástico.

Através da GSA, podemos relacionar os diferentes padrões presentes em toda ST. A Tabela 5.1 mostra os respectivos padrões de variabilidade para cada época.

Tabela 5.1 – padrões de variabilidade (1: Logístico; 2: Hénon; 3: Lorenz; 4: Estocástico com α=1,50; 5: Estocástico com α=1,67; 6: Estocástico com α=2,00) para as respectivas épocas (S: Seca; C: Cheia).

Ano	**S**	**C**	**Ano**	**S**	**C**	**Ano**	**S**	**C**	**Ano**	**S**	**C**	**Ano**	**S**	**C**
1900	5	2	1922	5	1	1944	4	2	1966	6	1	1988	6	1
1901	5	2	1923	5	1	1945	2	1	1967	5	2	1989	5	2
1902	4	2	1924	6	2	1946	6	2	1968	6	1	1990	6	1
1903	6	1	1925	5	4	1947	5	1	1969	5	4	1991	4	2
1904	6	1	1926	6	1	1948	4	2	1970	6	1	1992	4	4
1905	6	2	1927	1	2	1949	5	2	1971	6	1	1993	4	2
1906	2	2	1928	2	2	1950	6	2	1972	6	1	1994	2	2
1907	4	2	1929	4	1	1951	6	2	1973	6	2	1995	5	1
1908	5	1	1930	1	1	1952	6	1	1974	2	1	1996	6	2
1909	6	2	1931	6	1	1953	4	2	1975	4	1	1997	5	2
1910	5	2	1932	5	2	1954	5	1	1976	5	2	1998	4	2
1911	4	2	1933	4	2	1955	4	1	1977	6	2	1999	1	2
1912	6	4	1934	6	2	1956	5	1	1978	1	2	2000	4	1
1913	5	4	1935	5	1	1957	4	2	1979	1	1	2001	2	1
1914	5	2	1936	5	1	1958	6	2	1980	5	4	2002	6	2
1915	3	4	1937	2	1	1959	1	2	1981	1	1	2003	5	2
1916	4	2	1938	4	1	1960	5	1	1982	5	2	2004	6	4
1917	6	1	1939	5	2	1961	5	4	1983	1	2	2005	2	2
1918	4	1	1940	6	2	1962	2	1	1984	5	2	2006	6	2
1919	3	4	1941	5	4	1963	6	1	1985	6	3	2007	6	2
1920	6	4	1942	6	2	1964	5	1	1986	6	2			
1921	5	2	1943	1	2	1965	5	2	1987	5	1			

Um resultado relevante é que a GSA conseguiu classificar os padrões de cada época de seca e cheia com quantidade de dados tão reduzida. Em 100% das classificações, tanto no período de cheia quanto no de seca, a GSA não se confundiu entre os padrões. Esse resultado é significativo porque, mesmo lidando com séries muito curtas, a técnica da GSA conseguiu classificar separadamente os padrões. A GSA já tinha obtido êxito com outras séries bem curtas (DANTAS, 2008).

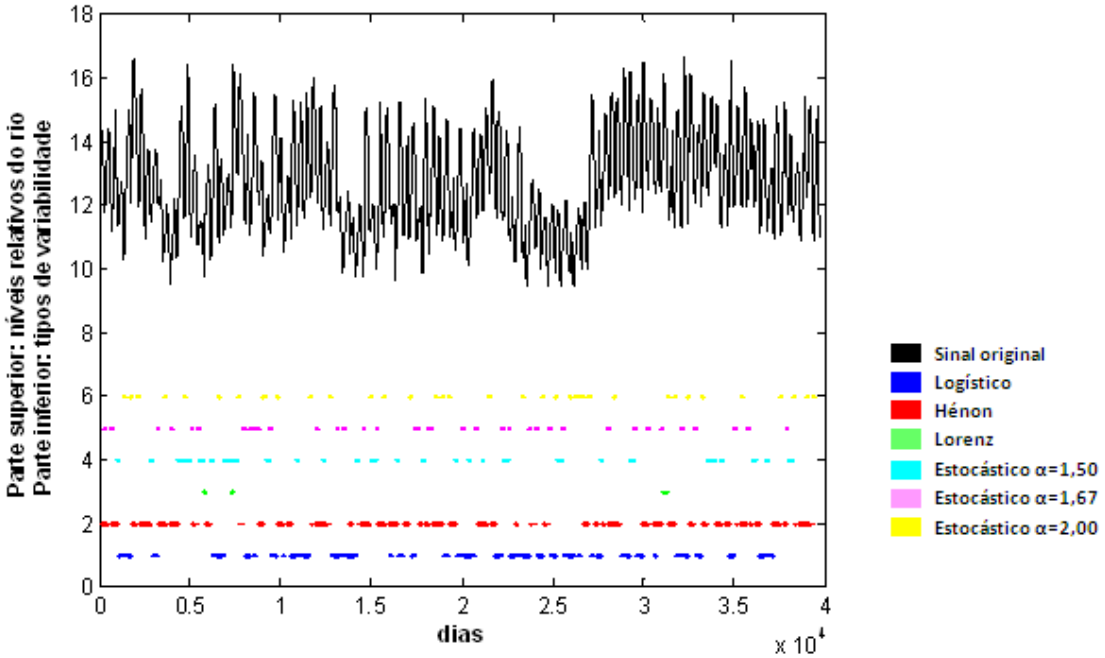


Figura 5.2 – ST completa (107 anos) de nível do rio Paraguai (parte superior), juntamente com os tipos de variabilidade para cada época (parte inferior).

Através da Figura 5.2, podemos visualizar a variação entre os padrões de variabilidade encontrados na ST do rio Paraguai (para ver com mais detalhe, veja Figura 5.3). A época de cheia é predominantemente marcada por variabilidade caótica, sendo que os subtipos de caos que mais aparecem são o Logístico e o Hénon. O subtipo Lorenz aparece somente em dois períodos de seca e em um período de

cheia. Já a época de seca é marcada principalmente por variabilidade estocástica sendo bem distribuído o aparecimento dos três subtipos estocásticos (α=1,50; α=1,67; α=2,00).

A Figura 5.3 corresponde à visualização da análise no decorrer dos três primeiros anos e serve para observarmos com maior precisão a classificação em cada período em separado.

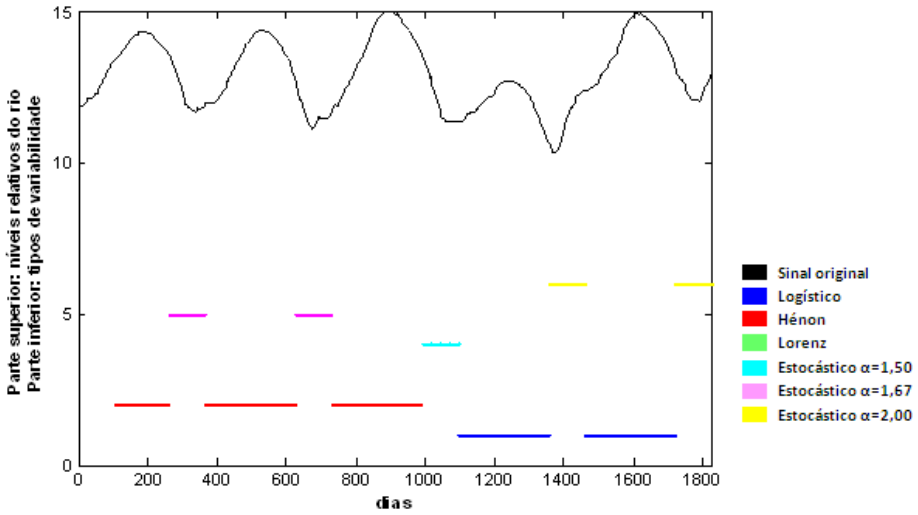
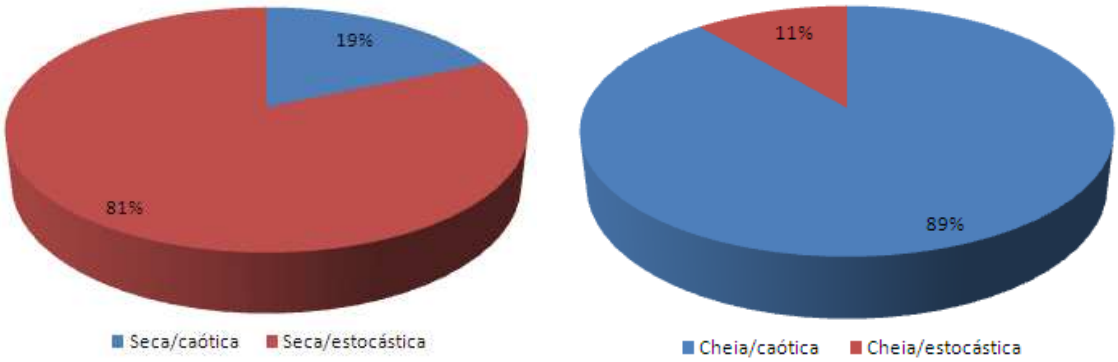


Figura 5.3 – ST parcial (3 anos) de nível do rio Paraguai (parte superior), juntamente com os tipos de variabilidade para cada época (inferior).



(a) (b)

Figura 5.4 – Distribuição de padrões caóticos e estocásticos nas épocas de (a) Seca e (b) Cheia.

A caracterização de padrão predominantemente caótico para a época de cheia e de padrão predominantemente estocástico para a época de seca era esperada devido às características observadas de persistência (época de cheia) e antipersistência (época de seca), segundo DANTAS (2008). Porém, observamos que algumas épocas de seca foram classificadas em alguns períodos como caóticas e épocas de cheia classificadas como estocásticas, conforme podemos visualizar na Figura 5.4.

6 CONCLUSÃO

Este trabalho estuda através da Análise Espectral Gradiente (GSA) a variação dos padrões de variabilidade no perfil temporal do nível do rio Paraguai ao longo de 107 anos (a partir de 1900), conjugando técnicas de Análise por Wavelets (Escalogramas e Multirresolução) com Análise de Padrões Gradientes. A GSA é uma metodologia inovadora que tem se mostrado robusta para analisar ST curtas.

Novamente a metodologia pôde ser testada em séries muito curtas, com tamanho da ordem de 105 pontos. Mesmo em tais condições a técnica mostrou-se eficiente para a classificação de processos relacionados às épocas de cheia e seca do rio com diferentes padrões de variabilidade relacionados a processos estocásticos e caóticos. Esses resultados são inéditos e devem repercutir de forma positiva na estratégia de pesquisa da Embrapa no Pantanal para os próximos anos.

Na seqüência dessa investigação, podemos estudar de maneira mais criteriosa as relações entre padrões de variabilidade subseqüentes, como por exemplo, a justificativa de quase toda seca classificada como padrão estocástico de α=1,5 ser seguida por uma cheia de padrão estocástico Hénon. Com isso, podemos sugerir a criação de um sistema de previsão do ciclo de cheia baseada na variabilidade temporal anual do ciclo de cheia do rio.

REFERÊNCIAS BIBLIOGRÁFICAS

ASSIREU, A. T. **Relatório de atividade de campo.** Período: 22/05/2005 a 25/05/2005. Disponível em: http://www.dpi.inpe.br/sima. Acesso em: 20 de maio de 2008.

BARBER, C. B.;DOBKIN, D. P.; HUHDANPAA, H.T. **ACM transactions on mathematical software**, v. 469, p. 22, 1996.

BARBOSA, E. B. M. **Estudo da variabilidade de séries temporais ambientais através de espectros de assimetria implementados em software científico livre.** 2006-05-26. 107 p. (INPE-14670-TDI/1223). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos. 2006.

BARLETT, K.B; HARRISS, R. C. Review and assesment of methane emissions from wetlands. **Chemosphere**, v. 26, 1993, p. 261-320.

BARONI, M. P. M. A.; CONCEIÇÃO, M. V.; ROSA, R. R.; PERSSON, C.; ARWIN, H.;SILVA Jr. E. F.; ROMAN, L. S.; NAKAMURA, O.; PEPE, I. ; SILVA, A. F. **J. of Non-Crystalline Solids**, v. 3734, p. 352, 2006a.

BARONI, M. P. M. A.; ROSA, R. R.; FERREIRA DA SILVA, A.; PEPE, I.; ROMAN, L. S.; RAMOS, F. M.; AHUJA, R.; PERSSON, C. ; VEJE, E. **Microelectronics Journal**, v. 290, p. 37, 2006b.

BRASIL. Ministério do Interior. **Descrição física e recursos naturais, estudo de desenvolvimento integrado da Bacia do Alto Paraguai** (Convênio Governo Brasileiro-UNDP-OEA) – Brasília: Superintendência do Desenvolvimento da Região Centro-Oeste, 1979. Relatório da primeira fase.

BUCHER, E.H.; BONETTO, A.; BOYLE, T.; CANEVARI, P.; CASTRO, G.; HUSZAR, P. ; STONE, T. **Hidrovia:** an initial environment examination of the Paraguay-Paraná Waterway – Wetland for the Americas, Massachusetts: Manomet, 1993. Publication Number 10.

DANTAS, M. S. **Análise espectral de padrões-gradiente de séries temporais curtas.** 2008. 157 p. (INPE-15676-TDI/1450). Dissertação (Mestrado em Computação Aplicada) - Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2009. Disponível em: <http://urlib.net/sid.inpe.br/mtc-m18@80/2009/02.05.10.55>. Acesso em: 26 set. 2009.

FLANDRIN, P. Fractional Brownian Motion and Wavelets. In: Farge, M.; Hunt, J.C.R.; Vassilicos, J.C., (eds). **Wavelets, fractals, and Fourier transforms**. Oxford: Claredon Press-Oxford, 1994, p.109-122.

FLANDRIN, P. Wavelet analysis and synthesis of fractional brownian motion. **IEEE Transactions on Information Theory**, v.38, n. 2, 1992, p.910-917.

HAMILTON, S. K.; SIPPEL, S. J.; MELACK, J. M. Inundation patterns in the Pantanal wetland of South America determined from passive microwave remote sensing. **Arch. Hydrobiol.**, v.137, n.1, 1996.

INSTITUTO NACIONAL DE PESQUEISAS ESPACIAIS (INPE). **Projeto balanço do carbono de Furnas Centrais Elétricas.** Período: 23/01/2005 a 11/02/2006. Disponível em: http://www.dpi.inpe.br/sima/bancos. Acesso em 20 de maio de 2008 (a).

JUNK, W. J. Wetlands of tropical South America. In: WHIGHAM, D., DYKY-JOVÁ, D. e HEJNÝ, S. (eds.). **Wetlands of the world**: inventory, ecology and management, vol I Kluwer, 1993. p. 679-739.

MALAMUD, B. D.; TURCOTE, D. L. Self-affine time series: measures of weak and strong persistence. **Journal of Statistical Planing and Inference**, v. 80, 1999, p. 173-196.

MALLAT, S. A. Theory for multiresolution signal decomposition: the wavelet representation, **IEEE Trans. Pattern Analysis and Machine Intelligence**, v. 11, n.7, 1989b, p.674-693.

MANDELBROT, B. B.; VAN NESS, J.W. Fractional brownian motions, fractional noises and applications. **SIAM Review**, v.10, n.4, 1968, p.422-437.

MORETTIN, P. A. **Ondas e ondaletas**: da análise de Fourier à análise de ondaletas. 1. Ed. São Paulo: Edusp, 1999.

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

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

RICHEY, J. E.; VICTORIA, R. L.; SALATI, E.; FORSBERG, B.R. The biogeochemistry of a major river system: the Amazon case study. – In: DEGENS, E. T.,

KEMPE, S. e RICHEY, J. E. (eds.). **Biogeochemistry of major world rivers**. New York: John Wiley, 1991.

ROSA, R. R.; BARONI, M. P. M. A.; ZANIBONI G. T.; FERREIRA DA SILVA, A.; ROMAN, L. S.; PONTES, J.; BOLZAN, M. J. A. **Physica A**, v. 666, p. 386, 2007.

ROSA, R.; CAMPOS, M.; RAMOS, F.; VIJAYKUMAR, N.; FUJIWARA, S.; SATO, T. Gradient Pattern Analysis of structural dynamics: application to molecular system relaxation. **Brazilian Journal of Physics**, v. 33, n. 3, p. 605–610, 2003.

ROSA, R.; PONTES, J.; CHRISTOV, C.; RAMOS, F.; NETO, C.; REMPEL, E.; WALGRAEF, D. Gradient Pattern Analysis of Swift-Hohenberg dynamics: phase disorder characterization. **Physica A**, v. 283, n. 1-2, p. 156–159, 2000.

ROSA, R.; SHARMA, A.; VALDIVIA, J. Characterization of asymmetric fragmentation patterns in spatially extended systems. **International Journal of Modern Physics B**, v. 10, n. 1, p. 147–163, 1999.

TEWFIK, A.H.; KIM, M. Correlation structure of the discrete wavelet coefficients of fractional brownian motion. **IEEE Transactions on Information Theory**, v.38, n.2, p.904-909, 1992.