

3 A Análise do Sinal

3.1. Introdução

Os modelos climatológicos mais modernos utilizam uma gama variada de ferramentas de análise baseadas em técnicas de processamento digital de sinais. Este capítulo descreve resumidamente alguns dos modelos mais conhecidos na literatura e as vantagens e limitações identificadas em cada um.

3.2. Filtros de Lanczos

O fenômeno climatológico focalizado é tratado como uma anomalia que ocorre com determinada periodicidade. O objetivo da utilização do Filtro de Lanczos [31] é ressaltar o comportamento de um sinal em determinadas frequências (ou periodicidades) em estudo.

O espectro $G(f)$ do sinal original $g(t)$ é multiplicado pela resposta em frequência do filtro $\Omega(f)$ determinado pela transformada de Fourier dos pesos dados pela expressão:

$$w(n) = \frac{\text{sen}(2pf_{c2}n)}{pn} - \frac{\text{sen}(2pf_{c1}n)}{pn} \quad (1)$$

onde f_{c2} e f_{c1} são as frequências de corte em unidades de tempo que definem a banda passante e n é o índice dos pesos.

Segundo Duchon [31], a quantidade mínima de pesos é determinada pela expressão:

$$n = \frac{1.3}{f_{c2} - f_{c1}} \quad (2)$$

A maior vantagem do filtro de Lanczos é a sua capacidade de ser facilmente sintonizado para a banda de interesse ajustando-se os parâmetros f_{c1} e f_{c2} mostrados nas equações. Para sinais como a *OLR*, que possuem amostragem

diária, uma maneira de filtrar anomalias na banda de oscilações com períodos entre 20 e 90 dias seria construir um filtro com uma fina banda passante, muito próxima da origem, com a maior porção do espectro entre $p/45$ e $p/12,5$, o que requereria construções com altíssimas ordens. Uma abordagem alternativa, e comumente adotada [21] em análises climatológicas, consiste em diminuir a resolução temporal do sinal calculando-se pântadas, médias não sobrepostas de intervalos de 5 dias do sinal. Este processo equivale a decimar o sinal de entrada [32], expandido o espectro original, fazendo com que a nova banda passante do filtro sintonizado para períodos entre 20 e 90 dias esteja entre $p/9$ e $p/2,5$. A Figura 9 ilustra o espectro em frequência de uma realização do Filtro de Lanczos com 75 elementos e banda passante ente $p/9$ e $p/2,5$. Fica claro, na figura, a pequena capacidade de rejeição (8dB) do filtro. Para facilitar a visualização do espectro respostas inferiores a -10db foram suprimidas do gráfico sem prejuízo da análise.

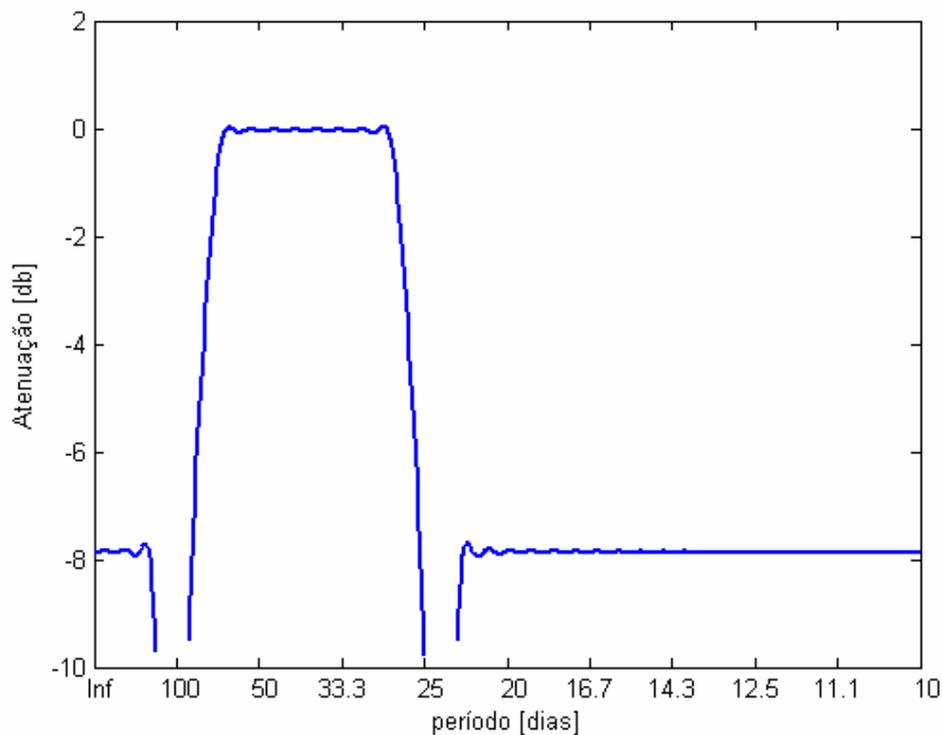


Figura 9 - Resposta em frequência do Filtro de Lanczos com 75 pesos, sintonizado para filtrar médias de 5 dias do sinal em uma banda passante de 90 a 25 dias.

É possível observar, na Figura 10, o resultado da filtragem de uma série de pântadas de *OLR* para a coordenada 5N, 60W, no ano de 1985, com o filtro

desenvolvido. Apenas o evento entre abril e julho fica levemente evidenciado, e mesmo assim, com pouca resolução temporal. Outros pequenos fenômenos no decorrer do ano não são detectados.

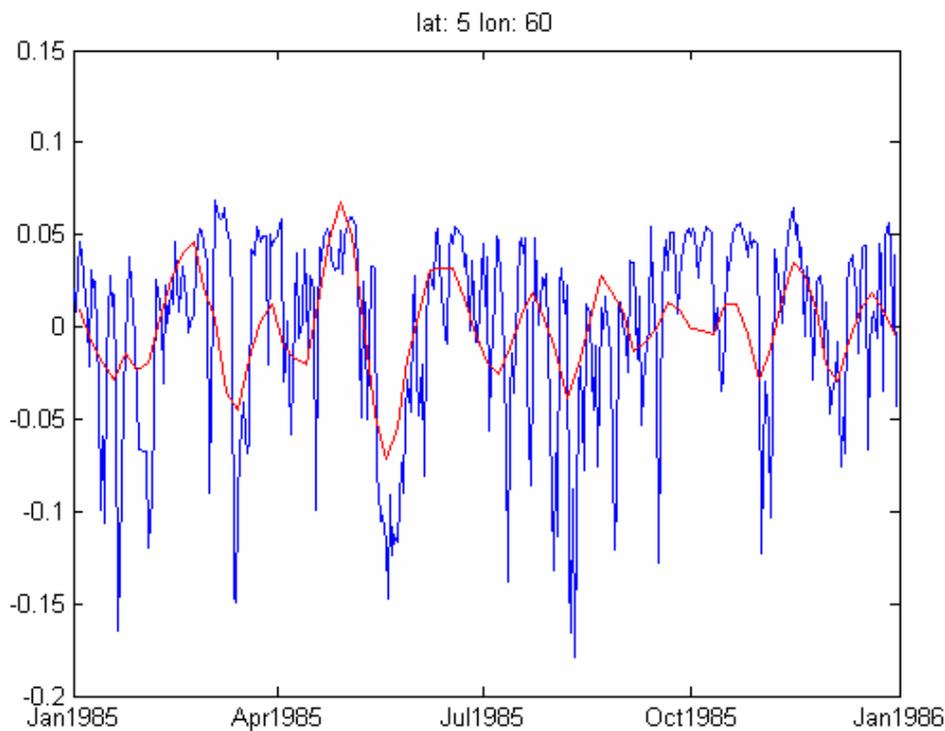


Figura 10 - Séries normalizadas: OLR na coordenada 5N, 60W, ano de 1985, (em azul), pêntadas filtradas da mesma série (em vermelho).

3.3. Wavelets

A decomposição por *wavelets* [32] é uma técnica bastante conhecida na área de processamento digital de sinais e vem sendo constantemente utilizada na análise de sinais climatológicos, mostrando-se bastante útil na identificação do comportamento periódico dos sinais, possibilitando uma modelagem da estrutura temporal e regional dos fenômenos climatológicos.

Resumidamente, tomando um sinal no tempo $g(t)$, a transformada wavelet $G(f, T)$ é dada por:

$$G(f, T) = \frac{1}{\sqrt{f}} \int_{-\infty}^{+\infty} \mathbf{w}\left(\frac{t-T}{f}\right) g(t) dt \quad (3)$$

onde

$w(t, T, f)$ é uma função conhecida como Wavelet-mãe, f é o fator de escala em frequência e T é fator de translação no tempo.

Há várias classes de funções-mãe [33-34]; a seleção deve ser feita levando em conta o tipo de resposta da função e o sinal em estudo.

A Figura 11 apresenta uma aplicação desta técnica à análise do impacto do fenômeno *Madden-Julian* ao sinal da *OLR* no Oceano junto ao Sul da China [35]. Pode-se apreciar o sinal mais característico das oscilações *MJO*: as periodicidades do sinal em torno de 30 a 60 dias calculados com médias de *OLRs* no oceano ao Sul da China, no ano de 1998. Uma observação mais detalhada da figura mostraria que estas componentes não são as únicas. Pode-se identificar também uma componente de periodicidade entre 10 e 20 dias, que aparece de forma mais fraca entre abril e maio, e de forma mais evidente de agosto a outubro, sendo levemente deslocada para a banda de 15 a 25 dias durante o mês de setembro. Em outras palavras, a padronização através da frequência 30 a 60 dias poderia ser insuficiente para uma caracterização precisa deste sinal.

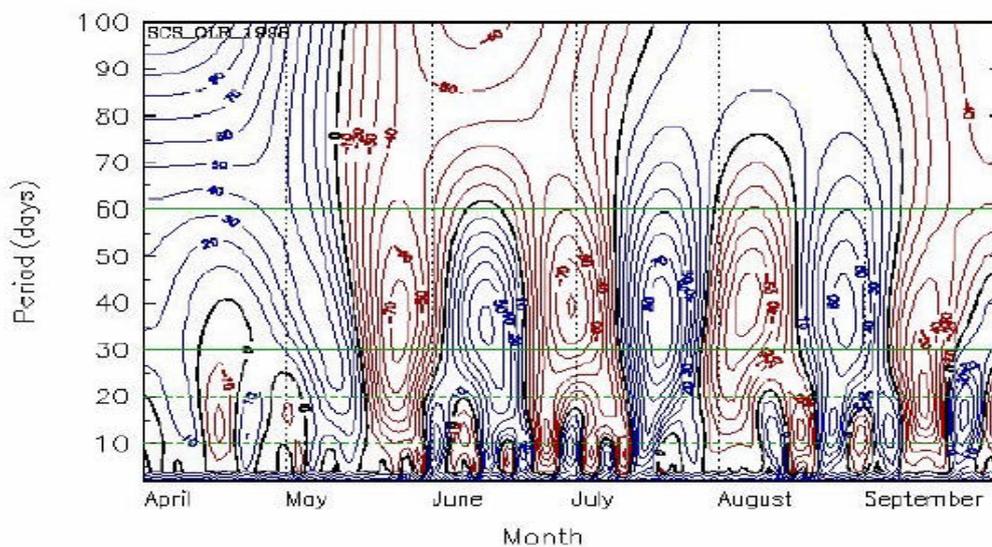


Figura 11 -Transformada Wavelet de um sinal de *OLR* no Oceano ao Sul da China [35]

As periodicidades apontadas pelas transformadas Wavelets podem ser utilizadas para a identificação de relações climatológicas – espectros semelhantes podem ser tomados como indicação de uma relação climatológica entre diferentes pontos ou regiões, o que é bastante útil quando o fenômeno é regular. A desvantagem surge quando a dinâmica das relações não é uniforme, como no caso

da propagação do evento *El Nino* entre os dois pontos apresentada na Figura 8, e na análise do oceano no Sul da China ilustrada na Figura 11. Neste caso, torna-se necessária uma análise ponto a ponto (ou região a região) e evento a evento – o que exigirá a aplicação de outras técnicas.

3.4. Funções Ortogonais Empíricas

O método das Funções Ortogonais Empíricas (*Empirical Orthogonal Functions*) – EOF, [27] corresponde basicamente a uma análise estatística que busca a maior correlação entre um conjunto de variáveis quaisquer. Este método tem sido exaustivamente utilizado na climatologia [17,18,29,36-38] para encontrar relações que possam descrever conexões ou dinâmicas de propagação entre distintas regiões.

O Método EOF parte da decomposição SVD de uma matriz X de dimensão $p \times n$, resultando em

$$X = U \Sigma V^t \quad (4)$$

onde U é uma matriz $p \times p$ constituída por autovetores ortonormais ditos à esquerda, S é uma matriz diagonal de dimensão $p \times n$ com os autovalores associados e V é uma matriz $n \times n$ constituída por autovetores ortonormais ditos à direita.

É imediato mostrar que

$$XX^t = U \Sigma V^t V \Sigma^t U^t = U \Sigma I_V \Sigma^t U^t = U I_U \quad (5)$$

enquanto que

$$X^t X = V \Sigma^t U^t U \Sigma V^t = V \Sigma^t I_U \Sigma V^t = V I_V \quad (6)$$

onde I_V é uma matriz Identidade de dimensão $n \times n$, enquanto que I_U é uma matriz Identidade de dimensão $p \times p$. λ corresponderá à porção associada aos elementos não nulos de $\Sigma \Sigma^t$ ou $\Sigma^t \Sigma$, respectivamente.

O Método EOF tem sido exaustivamente aplicado à análise de sinais climatológicos, buscando relações no tempo ou no espaço entre variáveis climatológicas. A referência [39] apresenta uma aplicação à busca de relações entre sinais de *OLRs* na região entre a China e a Austrália. O algoritmo aplicado pode ser descrito através dos seguintes passos:

- 1- Montagem da matriz X de dados – a partir das séries de anomalias, já filtradas no período 20 a 120 dias. A matriz X é composta pelos sinais de OLR distribuídos na região em análise, discretizada a cada 2.5° . A matriz tem dimensão $M \times N$, onde M é o número de quadrículas que formam a região e N é o índice temporal que referencia as séries de OLR s.

A Figura 12 ilustra a matriz X , onde cada linha corresponde a uma série temporal associada a cada quadrícula.

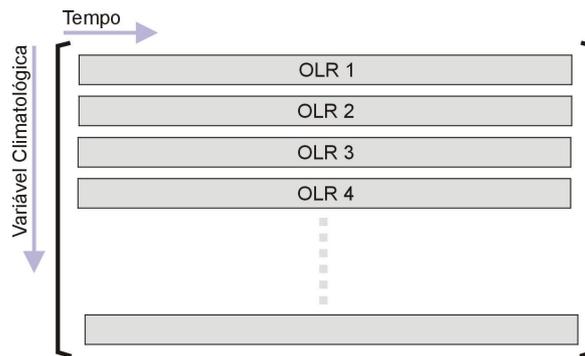


Figura 12 – Matriz X composta pelas diversas variáveis climatológicas (séries temporais de OLR associadas às diversas quadrículas)

- 2- Utilização das expressões (5) e (6) para cálculo de V , I e U . É interessante notar que a transformação

$$T_1 = XX^t \quad (7)$$

condensa a informação disponível no domínio do espaço, e portanto cada autovetor U_i encerra a distribuição espacial associada ao modo λ_i . Analogamente, a transformação

$$T_2 = X^t X \quad (8)$$

condensa a informação no domínio do tempo e portanto cada autovetor V_i encerra a distribuição temporal associada ao modo λ_i .

Os vetores V_i são conhecidos como as Componentes Principais (*Principal Components, PC*), e os vetores U_i constituem as Funções Ortogonais Empíricas (*Empirical Orthogonal Functions, EOF*). Matematicamente, os autovalores constituem os modos da matriz.: quanto mais significativos, maior a possibilidade da existência de um “moto” físico a ele associado. A escolha dos maiores autovalores e identificação dos seus motos físicos é, mais que um problema matemático, uma arte, que deve ser exercida por profissionais experientes, capazes de distinguir e interpretar os resultados oferecidos pelo método [38].

Por exemplo, no trabalho tomado como exemplo [39], foram identificados dois modos principais associados a suas respectivas componentes (21.11% e 20.38%, em uma base normalizada). São apresentadas as distribuições espaciais e temporais, evidenciando as regiões onde cada modo se manifesta mais intensamente e sua evolução temporal.

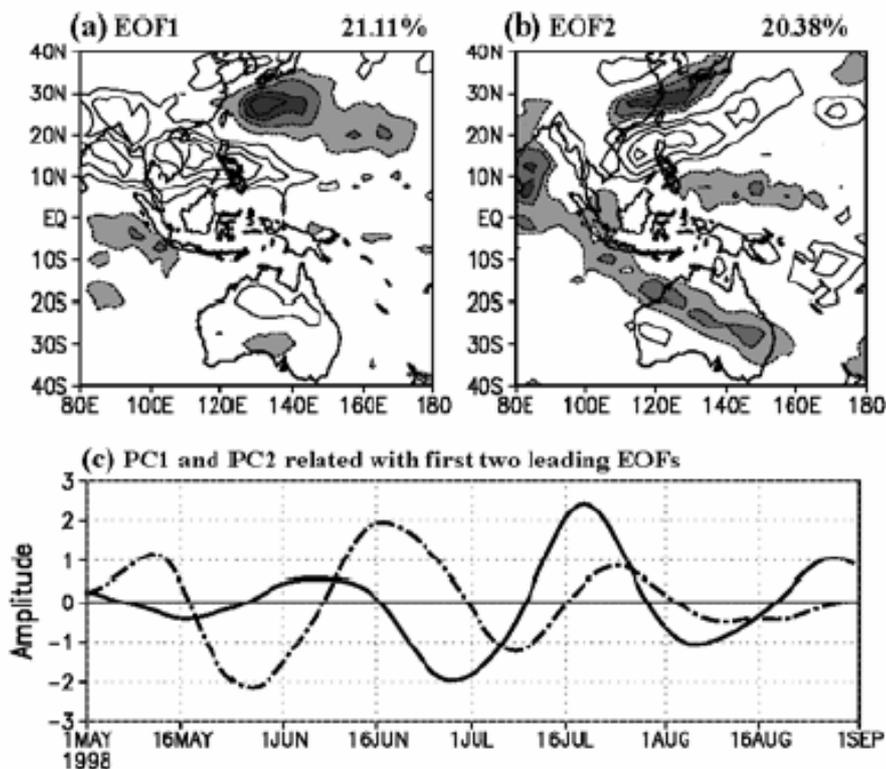


Figura 13 – Análise de EOF das *OLRs* – espacial e temporal – retirado da referência [39]

As vantagens do método são muitas: além da consistência e elegância matemática, é uma ferramenta que separa os modos de propagação e evidencia as inter-relações entre variáveis – inclusive pode ser aplicado a indicadores climáticos distintos, relacionando, por exemplo, temperaturas, ventos, pressão, etc. [40-42].

Detetamos, entretanto uma limitação: ao realizar correlações genéricas em um período de tempo qualquer – em princípio, quanto mais longo melhores seriam os resultados, parte-se do princípio que há uma estacionariedade da dinâmica climatológica. Em outras palavras, assume-se que os eventos repetem padrões comportamentais idênticos – o que nem sempre se verifica na prática, como ilustrado por Rui e Wang [20]. Por isso, a aplicação do método EOF poderá falhar na identificação de eventos anômalos, diferentes – justamente os mais difíceis de modelar e entender.

3.5. Análise Espectral Singular

A Análise Espectral Singular (*Singular Spectrum Analysis, SSA*) [43] pode ser vista como um caso particular do Método das Funções Ortogonais Empíricas, onde se deseja estudar o comportamento de uma única variável e conhecer suas características específicas, como tendências, sazonalidades, harmônicos etc. Este método pode ser utilizado com sucesso tanto na previsão futura quanto no preenchimento de lacunas em séries devidas à falta de informação.

A implementação da *SSA* parte da construção da chamada matriz trajetória formada por uma janela deslizante na série em estudo. Por exemplo, a matriz trajetória de uma série temporal formada a partir de uma janela de tamanho N é dada por

$$S = \begin{bmatrix} s_1 & s_2 & s_3 & \cdots & s_N \\ s_2 & s_3 & s_4 & \cdots & s_{N+1} \\ & & \vdots & & \\ s_N & s_{N+1} & s_{N+2} & \cdots & s_{2N-1} \end{bmatrix} \quad (9)$$

A aplicação das expressões (4), (7) e (8) a S explicitam os modos da matriz trajetória, detetando relações no tempo e realçando periodicidades e motos recorrentes.

A referência [44] apresenta a aplicação da *SSA* à análise dos eventos *El Nino* a partir de um dos índices mais utilizados para a sua identificação: o índice *SOI* (*Southern Oscillation Index*) [30], ilustrado na Figura 14 e definido como a diferença normalizada entre a pressão do ar no Pacífico Leste e Oeste. Em anos normais, este índice exhibe valores positivos; um evento de *El Nino*, entretanto, é associado à reversão nesta diferença de pressão e, conseqüentemente, a valores *SOI* negativos. A periodicidade de eventos *El Nino* não é clara: varia de três a cinco ou seis anos, o que dificulta a sua análise em termos recorrentes ou padronizados.

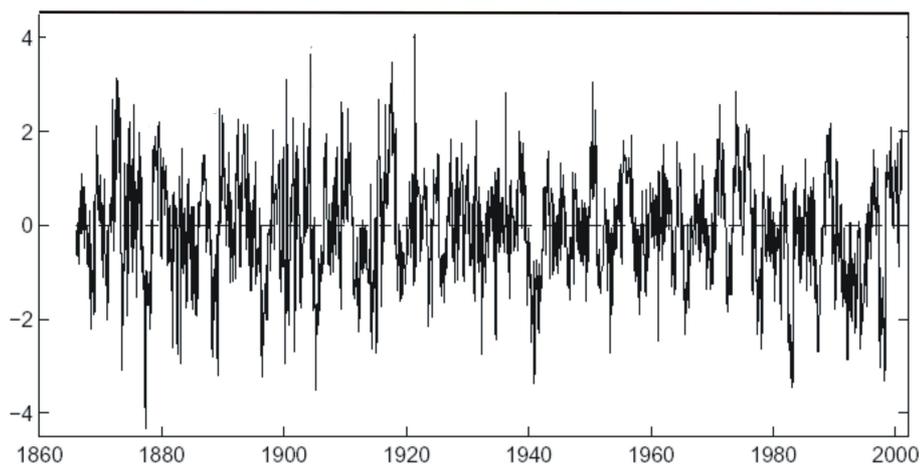


Figura 14 – Série Temporal do Índice SOI (Southern Oscillation Index) [44]

A Figura 15 apresenta as componentes principais para aplicação da SSA à série temporal SOI com uma janela de 72 meses. É possível observar recursividades com períodos que variam desde 20 a 60 meses.

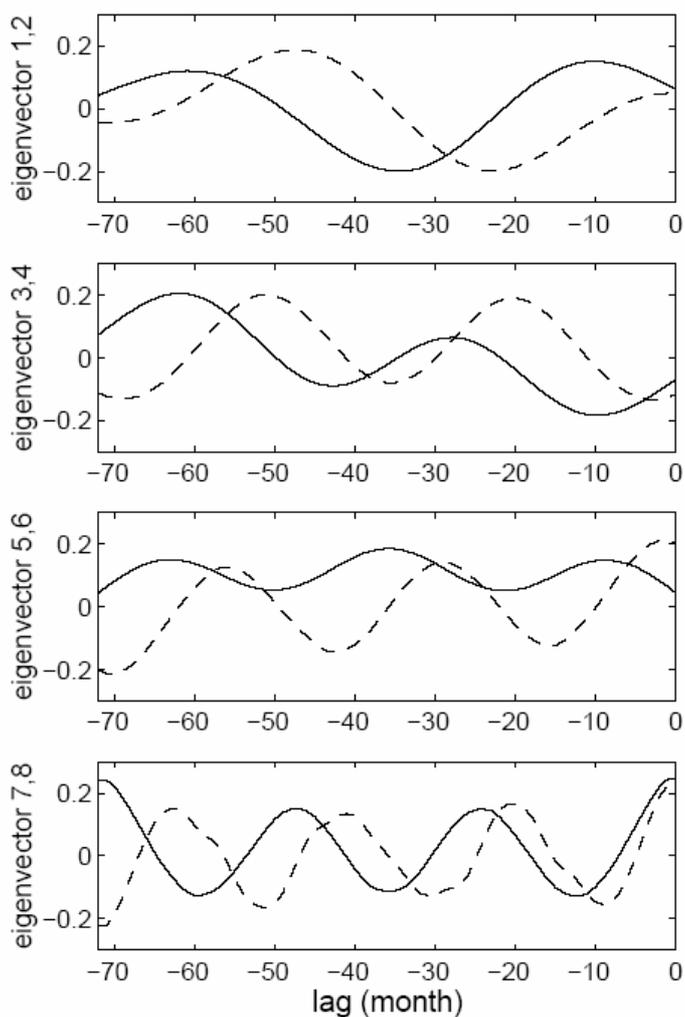


Figura 15 – Aplicação da SSA (*Singular Spectrum Analysis*) à série temporal do índice SOI [44].

3.6. Conclusões

Este capítulo apresentou os métodos e modelos mais utilizados na análise, identificação e previsão de variáveis climatológicas. Todos estes modelos, entretanto, buscam, por várias técnicas, identificar semelhanças, de modo a construir padrões de comportamento no espaço ou no tempo. Apesar de toda a sua potencialidade, o conceito da busca do semelhante pode ser insuficiente na detecção do *novo* – isto é, de fenômenos anômalos, diferentes do padrão, geralmente associados a eventos extremos e, por isso mesmo, imprevisíveis a partir das técnicas convencionais.

O próximo capítulo introduz a proposta deste trabalho: ao invés do semelhante, busca-se a identificação do *diferente*. O objetivo será desenvolver uma técnica de processamento de sinais que ressalte comportamentos anômalos no sinal sem a preocupação da padronização. Nosso objetivo será a identificação do início do evento, sua intensidade, persistência, duração e propagação.