

A Transformada *Wavelet* e sua Aplicação na Análise de Séries Hidrológicas

Celso Augusto Guimarães Santos¹; Paula Karenina de Macedo Machado Freire¹; Christopher Torrence²

celso@ct.ufpb.br; pulkymm@gmail.com; chris.torrence@exelisvis.com

Recebido: 31/05/12 - revisado: 20/09/12 - aceito: 25/03/13

RESUMO

Este artigo fornece um guia prático da transformada “wavelet” para a análise de dados hidrológicos. Para tanto, sua aplicação é feita a priori a uma série hipotética composta de quatro frequências para melhor ilustrar a técnica e em seguida é aplicada a duas séries pluviométricas de postos da bacia do rio São Francisco com dados do período de 1936 a 2011 com o objetivo de explorar esta técnica na análise de dados hidrológicos. A transformada “wavelet” corresponde a uma ferramenta que decompõe uma série unidimensional no duplo domínio tempo-frequência, permitindo assim a identificação dos principais modos da variabilidade, bem como a forma como estes variam no tempo. Ela é um avanço recente no processamento de sinais e baseia-se na convolução destes com sucessivas funções de diferentes escalas, as funções “wavelets”. Este trabalho faz uma explanação das características fundamentais das “wavelets” e exemplifica algumas de suas aplicações, usando-as em dados pluviométricos da bacia do rio São Francisco.

Palavras-chave: Ondaletas. Processamento de sinais. Séries temporais.

INTRODUÇÃO

A transformada *wavelet* é um avanço recente em processamento de sinais que tem atraído muita atenção desde seu desenvolvimento teórico em 1984 por Grossman e Morlet (1984). Seu uso tem aumentado rapidamente como uma alternativa à transformada Fourier (TF) em preservar os fenômenos locais, não-periódico e de multiescala. O termo *wavelet* vem do termo francês *ondelette*, significando onda pequena. Morettin (1999) traduziu-o para o português como ondaletas; entretanto, será adotado neste artigo o termo *wavelet*.

Esta análise de *wavelet* tem vantagem sobre as análises espectrais clássicas, porque permite analisar periodicidade de eventos em diferentes escalas da variabilidade temporal e não necessita de uma série estacionária. Assim, a ferramenta é apropriada para analisar eventos irregularmente distribuídos e séries temporais que contenham potências não-estacionárias em diferentes frequências. Por isso, a transformada *wavelet* está se tornando uma ferramenta comum para analisar variações locais de potência dentro da série.

Várias ciências aplicadas estão fazendo uso das *wavelets*, tais como astronomia, acústica, compressão de dados, engenharia nuclear, codificação de sub-bandas, processamento de imagens e sinais, neurofisiologia, música, ressonância magnética, reconhecimento de voz, ótica, fractais, radar, visão humana, matemática pura e geofísica, tais como convecção tropical, El Niño-Oscilação Sul, frentes atmosféricas frias, variabilidade de temperatura, dispersão de ondas oceânicas, crescimento e quebra de ondas, estruturas em fluxos turbulentos, e caracterização de vazões (BRAGA; SANTOS, 2010; FARGE, 1992; GRAPS, 1995; SANTOS; GALVÃO; TRIGO, 2003; SANTOS; IDEIÃO, 2006; SMITH; TURCOTTE; ISACKS, 1998; TORRENCE; COMPO, 1998).

A transformação de uma função (sinal) do domínio tempo para o domínio frequência pode ser feita através de diversas ferramentas matemáticas. Dentre elas, a mais conhecida é provavelmente a transformada de Fourier que foi desenvolvida em 1807 pelo matemático Joseph Fourier. A análise de Fourier baseia-se na afirmação de que toda função real periódica pode ser escrita como uma soma infinita de senos e cossenos, permitindo, desta forma, a identificação do sinal em um espaço unidimensional, neste caso na dimensão do tempo.

Porém, a maioria dos dados hidrológicos apresenta uma variabilidade aperiódica, o que torna a interpretação do espectro de potência de Fourier,

¹ - Departamento de Engenharia Civil e Ambiental/UFPB

² - Exelis Visual Information Solutions - Pearl East Circle - Boulder, CO, EUA

resultante da análise de sinais, complexa e passível de erros grosseiros em alguns casos.

Nesse contexto, surge a transformada *wavelet* como uma ferramenta recente e em pleno desenvolvimento que, em comparação com a transformada de Fourier, se mostra como uma alternativa bem mais robusta para a análise da frequência de sinais. No caso das *wavelets*, a decomposição do sinal não é mais feita em termos de senos e cossenos, como acontecia em Fourier, mas em termos de funções localizadas no tempo e sem escala fixa, fornecendo desta maneira não só a identificação da frequência como também a sua localização no tempo.

As seções seguintes descreverão a transformada *wavelet*, os dados hidrológicos e, por fim, uma aplicação da análise *wavelet* a tais dados, usando o programa desenvolvido por Torrence e Compo (1998).

A TRANSFORMADA WAVELET

Transformações matemáticas são aplicadas a sinais para obter mais informações sobre os mesmos, as quais não estão prontamente disponíveis no sinal original. Existem várias transformações que podem ser aplicadas, entre as quais as transformadas de Fourier são provavelmente as mais populares. Com o intuito de manter a localização da frequência no tempo em uma análise de sinal, uma possibilidade seria fazer uma transformada de Fourier janelada (*windowed fourier transform – WFT*), usando um certo tamanho de janela e deslizando-a ao longo do tempo, calculando a transformada rápida de Fourier (*fast fourier transform – FFT*) em cada tempo usando apenas os dados de dentro da janela. Isto resolveria o problema da localização da frequência, mas ainda dependeria do tamanho usado da janela. O principal problema com a *WFT* é o tratamento inconsistente de diferentes frequências: nas frequências baixas existem tão poucas oscilações dentro da janela que a localização da frequência é perdida, enquanto que nas frequências altas existem tantas oscilações que a localização do tempo é perdida. Finalmente, a *WFT* cai na suposição que o sinal pode ser decomposto em componentes senoidais, como já mencionado anteriormente.

Assim, para medir a estacionaridade de uma série temporal, é necessário calcular a variância contínua usando uma janela de tamanho fixo. Apesar da desvantagem de usar uma janela de tamanho fixo, a análise poderia ser repetida com uma variedade de tamanhos de janelas. Variando suavemente

o tamanho da janela, uma imagem das mudanças na variância *versus* tempo e tamanho da janela poderia ser obtida. O problema óbvio com esta técnica é a forma simples da função janela, que introduz efeitos de borda (oscilações). Usando tal janela, não haverá informação sobre o que está acontecendo dentro dela, mas teríamos apenas a energia média.

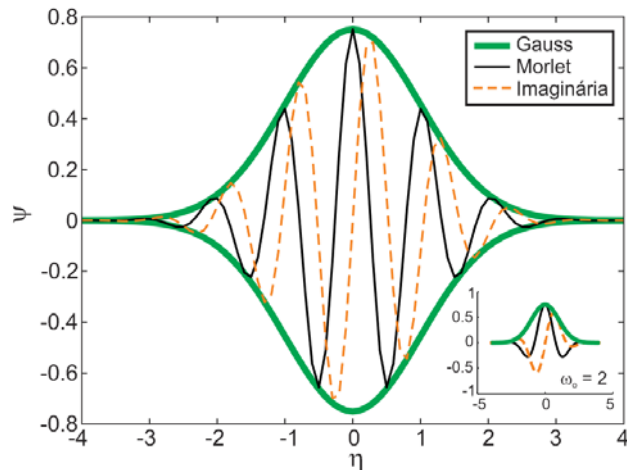


Figura 1 – A *wavelet*-mãe de Morlet

A análise *wavelet* objetiva em resolver estes problemas de decompor ou transformar uma série temporal unidimensional em uma imagem bidimensional difusa simultaneamente no domínio tempo-frequência. Assim, é possível obter informação sobre a amplitude de qualquer sinal periódico dentro da série, e como esta amplitude varia com o tempo.

Um exemplo de um “pacote” de onda, de duração finita e com uma frequência específica, é mostrada na figura 1. Tal formato poderia ser usado como uma função janela para a análise da variância. Esta pequena onda (*wavelet*) tem a vantagem de incorporar uma onda de período determinado, bem como ser finita na extensão. De fato, a *wavelet* mostrada na figura 1 (chamada de *wavelet* de Morlet) é nada mais que uma onda seno multiplicada por um envelope gaussiano (curvas superiores e inferiores em negrito).

Assumindo que a largura total desta *wavelet* seja cerca de 10 anos, é possível encontrar a correlação entre esta curva e os primeiros 10 anos da série. Este simples número dá uma medida (potência) da projeção deste pacote de onda sobre os dados durante estes 10 anos, i.e., quanto de [amplitude] o período de 10 anos assemelhasse a uma onda senoidal com esta largura [frequência]. Deslizando esta

wavelet sobre a série temporal, uma nova série temporal da amplitude de projeção *versus* o tempo pode ser construída.

Finalmente, a “escala” da *wavelet* pode ser variada mudando sua largura. Esta é a vantagem real da análise *wavelet* sobre espectro Fourier móvel. Para uma janela com certa largura, a *FFT* deslizante está ajustando diferentes números de ondas; i.e., pode haver muitas ondas de alta frequência dentro da janela, enquanto a mesma janela pode apenas conter poucas (ou menos que uma) ondas de baixa frequência. A análise *wavelet* sempre usa a *wavelet* com a mesma forma, apenas o tamanho da escala aumenta ou diminui com o tamanho da janela.

Além da amplitude de um sinal periódico, vale a pena obter informação sobre a fase. Na prática, a *wavelet* de Morlet mostrada na figura 1 é definida como um produto de uma onda exponencial complexa ($e^{i\omega_0\eta}$) e um envelope gaussiano ($\pi^{-0,25}e^{-0,5\eta^2}$):

$$\Psi_0(\eta) = \pi^{-0,25} e^{i\omega_0\eta} e^{-0,5\eta^2} \quad (1)$$

onde $\Psi_0(\eta)$ é o valor da *wavelet* no tempo adimensional η , e ω_0 é a frequência adimensional, igual a 6 neste estudo para satisfazer uma condição de admissibilidade; i.e., a função deve ter média zero e ser localizada no espaço do tempo e frequência para se “admissível” como uma *wavelet*. A linha tracejada representa a parte imaginária. No detalhe da figura 1 pode ser visto a *wavelet* de Morlet com $\omega_0 = 2$.

Esta é a função básica *wavelet*, mas será necessário, agora, algum artifício para mudar o tamanho total bem como deslizar toda a *wavelet* ao longo do tempo. Assim, as “*wavelets* escaladas” são definidas como:

$$\Psi\left[\frac{(n'-n)\delta t}{s}\right] = \left(\frac{\delta t}{s}\right)^{0,5} \Psi_0\left[\frac{(n'-n)\delta t}{s}\right] \quad (2)$$

onde s é o parâmetro de “dilatação” usado para mudar a escala, e n é o parâmetro de translação usado para deslizar no tempo. O fator $s^{-0,5}$ é uma normalização para manter a energia total da *wavelet* escalada constante.

Temos uma série temporal, com N valores de x_n , no índice de tempo n' . Cada valor é separado no tempo por um intervalo de tempo constante δt . A transformada *wavelet* $W_n(s)$ é apenas o produto interno (ou convolução) da função *wavelet* com a série temporal original:

$$W_n(s) = \sum_{n'=0}^{N-1} x_{n'} \Psi^* \left[\frac{(n'-n)\delta t}{s} \right] \quad (3)$$

onde o asterisco (*) denota conjugado complexo.

O somatório acima pode ser avaliado para vários valores da escala s (usualmente múltiplos da frequência mais baixa possível), bem como todos os valores de n entre as datas de início e fim. Uma figura bi-dimensional da variabilidade pode ser construída plotando a amplitude e a fase da *wavelet*. Então, uma série temporal pode ser, dessa forma, decomposta em um espaço tempo-fase da frequência usando uma *wavelet* (mãe) típica. O cálculo real da transformada *wavelet* pode ser feito pelo algoritmo seguinte (TORRENCE; COMPO, 1998): (a) escolhe-se uma *wavelet*-mãe; (b) encontra-se a TF da *wavelet*-mãe; (c) encontra-se a TF da série temporal; (d) escolhe-se uma escala mínima s_0 , e todas as outras escalas; (e) para cada escala, faz-se:

- I. Usando a Eq. (4), ou qualquer uma apropriada para a *wavelet*-mãe em uso, calcula-se a *wavelet*-filha em cada escala:

$$\Psi(s\omega_k) = \left(\frac{2\pi s}{\delta t}\right)^{0,5} \hat{\Psi}_0(s\omega_k) \quad (4)$$

onde o sinal \wedge indica a TF.

- II. Normaliza-se a *wavelet*-filha dividindo-a pela raiz-quadrada da variância total da *wavelet* (o total de Ψ^2 deve ser então um, preservando assim a variância da série temporal);
- III. Multiplica-se pela TF da série temporal;
- IV. Usando Eq. (5), inverte-se a transformada de volta para o espaço real;

$$W_n(s) = \sum_{k=0}^{N-1} \hat{x}_k \hat{\Psi}^*(s\omega_k) e^{i\omega_k n \delta t} \quad (5)$$

onde ω_k é a frequência angular, igual a $2\pi k/N\delta t$ para $k \leq N/2$ ou igual a $-2\pi k/N\delta t$ para $k > N/2$.

É possível calcular a transformada *wavelet* no domínio do tempo usando a Eq. (3). Entretanto, é mais simples usar o fato de que a transformada *wavelet* é a convolução entre as duas funções x e Ψ , e realizar a transformada *wavelet* no espaço Fourier usando a *FFT*; e (f) faz-se um mapa de contorno.

A figura 2 mostra um exemplo de *wavelets*-mãe que poderiam ser usadas, i.e., Paul e DOG (*Derivative of Gaussian*), a qual com $m = 2$ seria a *wavelet* chapéu mexicano (Figura 2a). Cada uma teria sua particularidade; e.g., a de Morlet tem uma maior precisão no domínio da frequência do que no tempo, enquanto que a de Paul tem uma maior precisão no domínio do tempo do que na frequência.

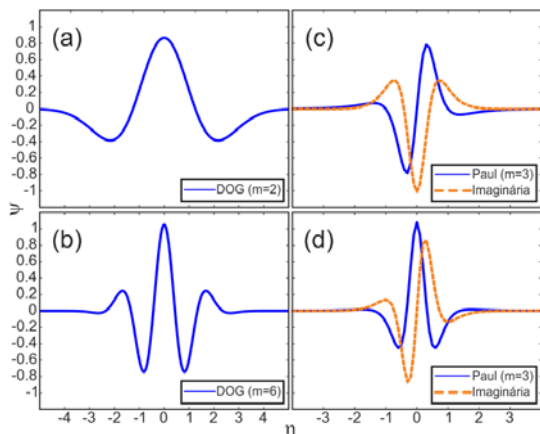


Figura 2 – Wavelet-mãe de (a) DOG com $m = 2$, (b) DOG com $m = 6$, (c) Paul com $m = 3$, e (d) Paul com $m = 4$

Como a transformada *wavelet* é um filtro passa-faixa (dispositivo que permite a passagem das frequências de uma faixa e elimina as frequências fora dessa faixa) com uma função de resposta conhecida (a função *wavelet*), é possível reconstruir a série original usando a deconvolução ou o filtro inverso. Isto é simples para a transformada *wavelet* ortogonal (que tem uma base ortogonal), mas para a transformada *wavelet* contínua é complicado pela

redundância no tempo e na escala. No entanto, esta redundância também torna possível reconstituir a série usando uma função *wavelet* completamente diferente. A função mais fácil seria a função delta δ , visto que trata-se apenas de um somatório (FARGE, 1992). Neste caso, a série reconstruída é apenas a soma da parte real da transformada *wavelet* sobre todas as escalas:

$$x_n = \frac{\delta_j \delta t^{0,5}}{C_\delta \psi_0(0)} \sum_{j=0}^J \frac{\Re\{W_n(s_j)\}}{s_j^{0,5}} \quad (6)$$

O fator $\psi_0(0)$ remove a escala de energia, enquanto $s_j^{0,5}$ converte a transformada *wavelet* em uma densidade de energia. O fator C_δ surge da reconstrução de uma função δ a partir da transformada *wavelet* usando a função $\psi_0(\eta)$. Este C_δ é uma constante para cada função *wavelet*. Note que se a série original for complexa, a soma do complexo $W_n(s)$ seria então utilizada.

EXEMPLOS DE SINAIS

Uma série composta por várias frequências (7,5, 30, 60 e 120 Hz) e o seu respectivo espectro de potência estão apresentados nas figuras 3a e 3b, respectivamente. A análise clássica de Fourier forneceria um espectro similar ao apresentado na figura 3c, a partir do qual se poderia saber que essas frequências (os picos na figura) formariam o sinal bruto, mas não se saberia quando as mesmas ocorreram.

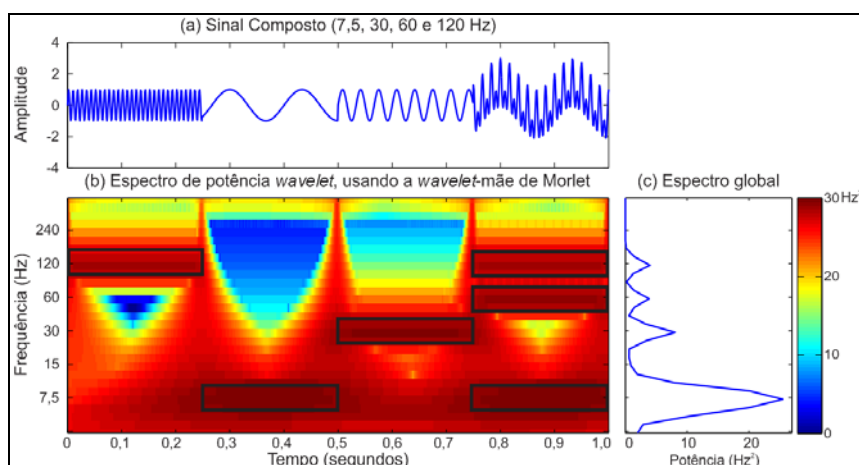


Figura 3 – (a) Sinal composto com as frequências de 7,5, 30, 60 e 120 Hz, (b) Espectro de potência wavelet, e (c) Espectro de potência global wavelet

Tabela 1 – Estatísticas descritivas das precipitações totais mensais dos postos São Francisco – SF (sub-bacia 44) para o período 1938 a 2005 e Delmiro Gouveia – DG (sub-bacia 49) para o período 1936 a 2011

Estatísticas	SF	DG	Unidades
Média	87,4	41,1	mm
Moda	0,0	0,0	mm
Mediana	38,4	26,8	mm
Amplitude	721,5	338,0	mm
Primeiro Quartil	1,1	5,0	mm
Terceiro Quartil	138,8	58,5	mm
Amplitude Inter-Quartis	137,7	53,5	mm
Desvio Abs. Médio	87,2	34,8	mm
Variância	12.714,4	2.302,2	mm ²
Desvio Padrão	112,8	48,0	mm
Coefficiente de Variação	1,3	1,2	—
Coefficiente de Assimetria	1,8	2,1	—
Coefficiente de Curtose	3,9	6,2	—
Excesso de Curtose	0,9	3,2	—

Já o espectro de potência *wavelet* (Figura 3b) fornece as informações de quando as referidas frequências ocorreram; i.e., 7,5 Hz ocorreu de 0,25 a 0,50 s e a partir de 0,75 s, 30 Hz de 0,50 a 0,75 s, 60 Hz a partir de 0,75 s, 120 Hz até 0,25 s, e depois a partir de 0,75 s. Note que até os primeiros 75 s, ocorreram três tipos de frequência (7,5, 30 e 120 Hz) em momentos diferentes, mas a partir de 75 s, as frequências de 7,5, 60 e 120 Hz ocorreram concomitantemente. Esta figura seria uma espécie de mapa topográfico, onde no eixo z estaria plotado a potência, onde as regiões mais altas identificam as frequências existentes. No presente caso, elas estão destacadas por um retângulo em negrito.

ANÁLISE DE DADOS PLUVIOMÉTRICOS

A análise *wavelet* foi escolhida para análise de dados hidrológicos, além das outras razões e vantagens descritas aqui, porque aplicações do tipo da análise padrão da Transformada Fourier a uma série temporal deve ser testada quando a série temporal cumpre duas importantes características, a saber: (1) estacionaridade; i.e., que não haja mudanças na média, variância, etc. ao longo da série temporal; e (2) que a série temporal possa ser descrita como uma soma de diferentes componentes periódicos (representada por simples funções harmônicas) para todo o período. Entretanto, a maioria das séries temporais da meteorologia e hidrologia não satisfaz ambos os requisitos. Na verdade, as séries temporais das ciências da Terra são geralmente

não-estacionárias e apresentam tendências ao valor médio, mudanças na variabilidade para determinados períodos. Além disso, muitas séries temporais hidrológicas, tais como precipitação, apresentam eventos irregularmente distribuídos com potência não-estacionária ao longo de várias frequências. Assim, sua estrutura intrínseca temporal não é bem representada pela superposição de uma frequência de poucos componentes como a derivada de uma análise de Fourier clássica.

Foram selecionados dois postos pluviométricos da bacia do rio São Francisco, i.e., o posto São Francisco (código 1544012) da sub-bacia 44 e o posto Delmiro Gouveia (código 937013) da sub-bacia 49, cujas séries compreendem respectivamente os períodos de 1938–2005 (Figura 4a) e 1936–2011 (Figura 5a). A tabela 1 mostra as estatísticas descritivas das séries pluviométricas a serem usadas na análise *wavelet*.

(1) Espectro de potência *wavelet*

Como os dados selecionados para esta aplicação são distribuídos mensalmente, os parâmetros para a análise *wavelet* foram configurados como $\delta t = 1$ mês e $s_0 = 2$ meses porque $s = 2\delta t$, $\delta_j = 0,25$ para fazer quatro sub-oitavas por oitava, e $j_1 = 8/\delta_j$ para fazer oito potência-de-dois com δ_j sub-oitavas cada.

A figura 4b mostra a potência (valor absoluto ao quadrado) da transformada *wavelet* para os totais pluviométricos mensais no posto São Francisco, enquanto que a potência da transformada *wavelet* para os totais mensais do posto Delmiro Gouveia são mostrados na figura 5b. Como dito anteriormente, o (valor absoluto)² fornece informação sobre a potência relativa em uma certa escala e em um certo tempo. Estas figuras mostram as oscilações reais das *wavelets* individuais, ao invés de apenas a sua magnitude. Observando a figura 4b, é claro que há mais concentração de potência entre a banda 8–16 meses, o que mostra que essa série temporal tem um sinal anual forte. Análises estatísticas clássicas aplicadas por autores anteriores para esta área têm mencionado a existência de importantes picos de baixa frequência. Aqui, é mostrado que tais resultados são enganosos já que picos significativos não foram alcançados para períodos de baixa frequência (e.g., 16 a 512 meses). Entretanto, o espectro de potência *wavelet* para os episódios de precipitação com escala característica de 4–8 meses apresenta um pico importante, quase significativo ao nível de 5%. A potência média na banda 8–16 meses também mostra os anos secos e úmidos; i.e., quando a potên-

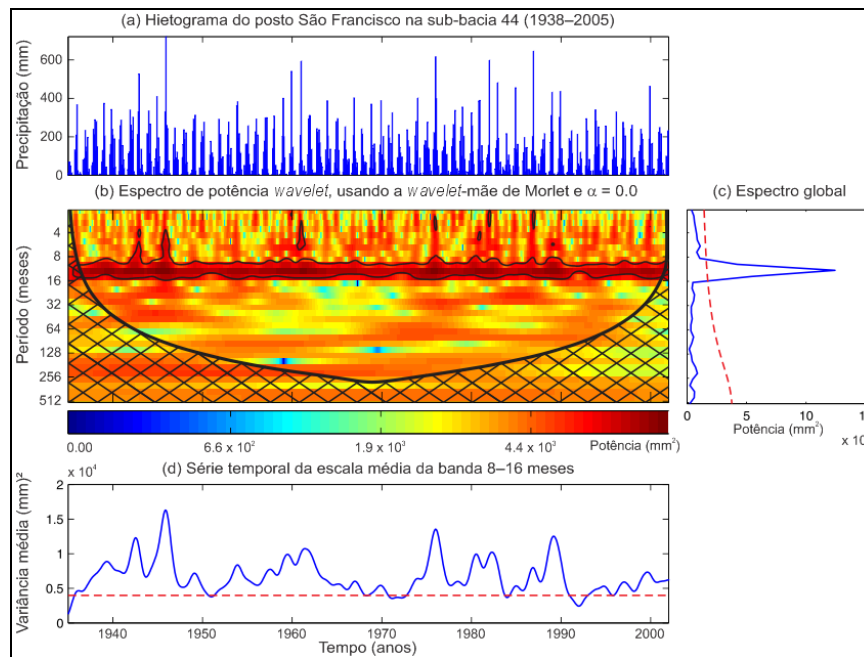


Figura 4 – (a) Hietograma do posto São Francisco (1544012). (b) Espectro de potência *wavelet*. A região hachurada é o cone de influência, onde o preenchimento com zero reduziu a variância. O contorno em negrito é o nível de 5% de significância, usando o ruído-branco ($\alpha = 0,0$) como espectro de fundo. (c) Espectro de potência global (linha azul). Linha tracejada é o nível de significância de 5% para o espectro global. (d) Série temporal da escala média da banda 8–16 meses. A linha tracejada é o nível de 95% de confiança assumindo o ruído-branco

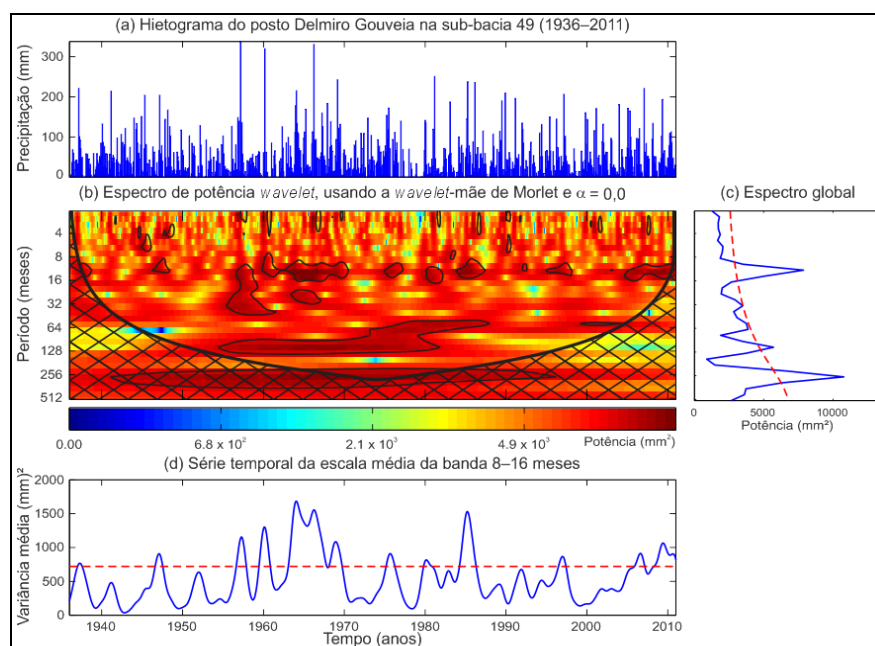


Figura 5 – (a) Hietograma do posto Delmiro Gouveia (937013). (b) Espectro de potência *wavelet*. A região hachurada é o cone de influência, onde o preenchimento com zero reduziu a variância. O contorno em negrito é o nível de 5% de significância, usando o ruído-branco ($\alpha = 0,0$) como espectro de fundo. (c) Espectro de potência global (linha azul). Linha tracejada é o nível de 5% de significância para o espectro global. (d) Série temporal da escala média da banda 8–16 meses. A linha tracejada é o nível de 95% de confiança assumindo o ruído-branco

cia diminui consideravelmente nesta banda, isto significa um ano seco e quando a potência é máxima significa um ano úmido. Por exemplo, na primeira metade da década de 90 os índices pluviométricos começaram a baixar bastante, o que provocou uma redução da potência nesta banda.

A figura 5b apresenta picos isolados, mas apesar das baixas de potência ao longo da banda 8–16 meses, a frequência anual (periodicidade de 12 meses) está presente também nesta série pluviométrica, podendo ser confirmada pela figura 5c.

A área hachurada nas Figuras 4b e 5b é o cone de influência, onde o preenchimento com zeros reduziu a variância. Como estamos lidando com séries temporais finitas, erros ocorrerão no início e no fim do espectro de potência *wavelet*. Uma solução seria preencher o final da série com zeros antes de aplicar a transformada *wavelet* e depois removê-los. Aqui a série temporal foi preenchida com zeros o suficiente para levar o seu comprimento N até a próxima potência de 2, limitando o efeito de bordas e aumentando a velocidade do cálculo da transformada Fourier. Este preenchimento com zeros introduz descontinuidades no final e diminui a amplitude próxima às bordas à medida que vai para as escalas maiores, já que mais zeros entram na análise. O cone de influência é a região do espectro de potência da *wavelet* na qual o efeito de bordas se torna importante e é definido como o tempo de envelopamento (*e-folding time*) para a autocorrelação da potência da *wavelet* em cada escala. Os picos nestas regiões foram presumidamente reduzidos em sua magnitude devido ao preenchimento com zeros. Assim, não é claro se a diminuição em qualquer banda de potência nesta área hachurada é uma diminuição real na variância ou um artefato do preenchimento. Para *wavelets*-mãe do tipo da do chapéu mexicano, seus cones de influência serão bem menores e assim será menos afetado pelos efeitos de bordas. Note também que para séries cíclicas, não há necessidade de preenchê-las com zero, e por isso não haverá cone de influência (Figura 3a).

O contorno em negrito nas mesmas figuras é o nível de significância de 5%, usando um espectro de fundo do tipo ruído-vermelho ou ruído-branco. Muitas séries temporais geofísicas podem ser modeladas como ruído-vermelho ou ruído-branco. Um modelo simples para ruído vermelho é um processo autoregressivo univariado lag-1. O lag-1 é a correlação entre a série temporal e ela mesma, mas deslocada (*lagged*) de uma unidade de tempo (BOX; JENKINS, 1976). No presente caso, seria um deslocamento de um mês. O lag-1 mede a persistência de uma anomalia de um mês para o próximo.

Quando o lag-1 está bem acima de 0,4, é recomendado calcular o lag-1 verdadeiro α , o qual pode ser calculado através de uma aproximação usando $\alpha = 0,5(\alpha_1 + \alpha_2^{0,5})$, onde α_1 é a autocorrelação lag-1 e α_2 é a autocorrelação lag-2, que é o mesmo que lag-1 sendo que deslocado de dois pontos ao invés de 1. Quando α_1 não estiver bem acima de 0,4, é recomendado modelar a série como ruído-branco ($\alpha = 0,0$). Como as séries analisadas dos postos São Francisco e Delmiro Gouveia mostram respectivamente um $\alpha_1 = 0,4253$ e $\alpha_2 = 0,1726$, e $\alpha_1 = 0,1818$ e $\alpha_2 = 0,1448$, os lag-1 verdadeiros α foram assumidos como sendo 0,0 para ambos os casos.

A hipótese nula é definida para o espectro de potência da *wavelet* como assumindo que a série temporal tem um espectro de potência médio; se um pico no espectro de potência da *wavelet* estiver significativamente acima deste espectro de fundo, então pode ser assumido que seja uma característica verdadeira com certo percentual de confiança. Por definição, “nível de significância de 5%” é equivalente a “nível de confiança de 95%”, e implica em um teste contra certo nível de fundo, enquanto o “intervalo de confiança de 95%” refere à abrangência de confiança em cima do valor dado. A confiança de 95% implica que 5% da potência da *wavelet* deve estar acima deste nível.



Figura 6 – Mapa da bacia do rio São Francisco com a localização das capitais próximas, de alguns municípios e dos dois postos pluviométricos analisados

(2) Espectro de potência global

A frequência anual (periodicidade de 12 meses) destas séries temporais é confirmada por uma integração da potência sobre o tempo, i.e., uma média temporal das potências existentes entre o início e fim da série analisada (Figuras 4c e 5c), que mostram nesta frequência um pico significativo acima do nível de confiança de 95% no espectro de potência global, assumindo ruído-branco, i.e., $\alpha = 0,0$ (Figuras 4c e 5c), representado pelas linhas tracejadas. Entretanto, a Figura 4c também apresenta um pico quase significativo (no nível de 5%) centrado na banda 4–8 meses. De fato, a maioria dos valores mensais extremos de precipitação no posto de São Francisco (valores acima de 300 mm na Figura 4a) corresponde a pulsos de potência significativa alta na banda 4–8 meses (Figura 4b). A série pluviométrica do posto Delmiro Gouveia é mais complexa que a do posto São Francisco. A série do posto Delmiro Gouveia apresenta além do sinal anual, mais dois sinais: 10 anos e 20 anos significativos ao nível de 5% (Figura 5c). O sinal decenal ocorre entre 1955 e 1985, enquanto que o sinal bidecenal esteve presente em quase toda série, mais precisamente, entre os anos de 1940 e 1995.

Este espectro de potência global fornece uma estimativa não tendenciosa e consistente do espectro de potência verdadeiro da série temporal, e assim ele é um meio simples e robusto para caracterizar a variabilidade das séries. Os espectros de potência globais podem ser usados para descrever a variabilidade da precipitação em hietogramas não-estacionários. Para regiões que não mostrem mudanças a longo prazo nas estruturas dos hietogramas, espectros de potência globais da *wavelet* são úteis para resumir a variabilidade temporal da região e compará-la com a precipitação em outras regiões. O formato do espectro de potência global é controlado primariamente pela distribuição das escalas características e, dessa forma, poderia ser considerado como um diagnóstico do regime hidroclimático, mesmo se tratando de uma ampla gama de tamanhos de bacias hidrográficas, porque uma diferença qualitativa clara poderia ser encontrada no espectro global dos hietogramas de regiões climáticas diferentes. Por exemplo, as duas séries pluviométricas apresentadas aqui possuem um sinal anual bem caracterizado, o que foi confirmado pelas figuras 4b, 4c, 5b e 5c; entretanto, os seus espectros globais possuem formas distintas, o que poderia ser usado na identificação de zonas hidrológicamente semelhantes, como já utilizado por Santos, Morais e Silva (2009). Pela figura 6, é possível confirmar que,

apesar de pertencerem a mesma bacia hidrográfica, elas pertencem a zonas com regimes pluviométricos distintos.

(3) Séries temporal da escala média

A potência da escala média da *wavelet* (Figuras 4d e 5d) é uma série temporal da variância média em certa banda, neste caso na banda 8–16 meses, usada para examinar a modulação de uma série por outra, ou modulação de uma frequência por uma outra dentro da mesma série temporal (SANTOS et al., 2001). A banda 8–16 meses foi escolhida, neste caso, para permitir a análise da variação pluviométrica anual (12 meses). As figuras 4d e 5d são feitas pelas médias das figuras 4b e 5b, respectivamente, sobre todas as escalas entre os meses 8 e 16, o que dá uma medida da variância média anual *versus* o tempo. Por exemplo, para a série pluviométrica do posto São Francisco (Figura 4d), o gráfico da variância mostra períodos distintos quando a variância de precipitação mensal foi baixa, e.g., a segunda metade da década de 30, início da década de 50, segunda metade da década de 60 até a primeira metade da década de 70, segunda metade dos anos 80 e anos 90, e picos importantes do espectro de potência podem ser identificados em 1946, 1976 e 1986, indicando claramente períodos mais úmidos que o normal. Analogamente para a série pluviométrica do posto Delmiro Gouveia (Figura 5d), destacam-se vários períodos secos, e.g., década de 40, 50 e 90 e períodos úmidos, década de 60 e 80, podem também ser identificados. As linhas tracejadas correspondem ao nível de 95% de confiança considerando também o ruído-branco.

(4) Filtragem

As séries apresentadas nas figuras 4d e 5d dão uma ideia de uma série filtrada (sem ruído), a qual poderia ser usada como entrada em modelos hidrológicos do tipo chuva-vazão ou em modelos hidrossedimentológicos, e.g., Santos et al. (2003), ao invés da série original, evitando que os ruídos entrassem na análise e na geração dos valores previstos, já que os mesmos não são informações verdadeiras do sinal bruto (original).

Entretanto, estas séries são apenas a variância ou potência, mas a análise *wavelet* pode também ser usada para filtrar o sinal bruto (TORRENCE; COMPO, 1998). Como já apresentado anteriormente, a transformada *wavelet* (Equação 5) é na verdade um filtro passa-faixa de forma uniforme, e de localização e largura variadas. Somando-se ao longo de

um subconjunto das escalas (j_1 até j_2), pode-se construir uma série *wavelet* filtrada:

$$x'_n = \frac{\delta j \delta t^{1/2}}{C_\delta \Psi_0(0)} \sum_{j=j_1}^{j_2} \frac{\Re\{W_n(s_j)\}}{s_j^{1/2}} \quad (7)$$

Este filtro tem uma função resposta dada pela soma das funções *wavelet* entre as escalas j_1 e j_2 . Esta filtragem pode também ser feita simultaneamente na escala e no tempo através da definição de um limiar de potência *wavelet*. Esta “eliminação de ruído” remove quaisquer regiões de baixa amplitude da transformada *wavelet*, que são presumivelmente devido ao ruído. Esta técnica tem a vantagem sobre as filtrações tradicionais à medida que remove ruído em todas as frequências, e pode ser usada para isolar eventos únicos que têm um espectro de potência amplo ou vários eventos que têm frequência variável. Uma descrição mais completa, incluindo exemplos é dada em Donoho e Johnstone (1994).

CONCLUSÃO

A transformada *wavelet* tem se mostrado como uma poderosa ferramenta para a análise de séries e, portanto, adequada para a análise de dados hidrológicos. Dessa forma, este artigo procurou fornecer um guia prático para o uso da transformada *wavelet* na análise de séries hidrológicas com exemplos ilustrativos e práticos.

Duas séries pluviométricas mensais com características distintas, postos São Francisco, sub-bacia 44, e Delmiro Gouveia, sub-bacia 49, da bacia hidrográfica do rio São Francisco, foram utilizadas para exemplificar a aplicação da análise *wavelet*. Os espectros de potência de ambas as séries mostraram uma grande concentração de potência na banda 8–16 meses, revelando uma periodicidade anual, o que foi confirmado pelos picos da integração dos vetores de magnitude da transformada sobre o tempo, mostrando novamente um forte sinal anual para um nível de confiança de 95%. A série pluviométrica do posto São Francisco pode ser considerada uma série simples, por apresentar apenas uma frequência anual; entretanto, a de Delmiro Gouveia mostrou-se mais complexa, apresentando três frequências significativas (1, 10 e 20 anos), para o mesmo nível de confiança, além de duas quase significativas (2,5 e 5 anos). A variância média no tempo dessas séries foi analisada através das séries da escala média entre 8–

16 meses, o que se mostrou útil na identificação de períodos secos e chuvosos (Figuras 4d e 5d).

Finalmente, estudos futuros poderão ser realizados com a análise *wavelet* da precipitação na bacia hidrográfica do rio São Francisco junto com a análise das séries fluviométricas com o objetivo de beneficiar modelos hidrológicos através de uma decomposição das séries originais e de sua filtragem visto que a série filtrada (sem ruído) poderá ser usada como entrada nos modelos hidrológicos, ao invés da série original, evitando com isso que os ruídos entrem na análise e na geração dos valores previstos, já que os mesmos não são informações verdadeiras da série original.

AGRADECIMENTOS

Os autores agradecem à Agência Nacional de Águas (ANA) pela disponibilização dos dados através da hidroweb, ao CT-Hidro/MCT e CNPq pelo apoio financeiro e ao Prof. Dr. Richard de Marques da UFPB pela colaboração na confecção do mapa da bacia hidrográfica do rio São Francisco.

REFERÊNCIAS

- BOX, G. E. P.; JENKINS, G. *Time series analysis: forecasting and control*. San Francisco: Holden-Day, 1976.
- BRAGA, I. Y. L. G.; SANTOS, C. A. G.: Viability of rainwater use in condominiums based on the precipitation frequency for reservoir sizing analysis. *Journal of Urban and Environmental Engineering*, v. 4, n. 1, p. 23-28, 2010. doi: 10.4090/juee.2010.v4n1.023028.
- DONOHOO, D. L.; JOHNSTONE, J. M. Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, v. 81, n. 3, p. 425-455, 1994. doi: 10.1093/biomet/81.3.425.
- FARGE, M. Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics*, v. 24, p. 395-457, 1992. doi: 10.1146/annurev.fl.24.010192.002143.
- GRAPS, A. An introduction to wavelets. *Computational Science and Engineering, IEEE*, v. 2, n. 2, p. 50-61, 1995. doi: 10.1109/99.388960.

GROSSMAN, A.; MORLET, J. Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM Journal on Mathematical Analysis*, v. 15, n. 4, p. 723-736, 1984. doi: 10.1137/0515056.

MORETTIN, P. A. *Ondas e ondaletas: da análise de Fourier à análise de ondaletas*. São Paulo: Editora da Universidade de São Paulo, 1999.

SANTOS, C. A. G.; GALVÃO, C. O.; SUZUKI, K.; TRIGO, R. M. Matsuyama city rainfall data analysis using wavelet transform. *Annual Journal of Hydraulic Engineering JSCE*, v. 45, p. 211-216, 2001.

SANTOS, C. A. G., GALVÃO, C. O.; TRIGO, R. M. Rainfall data analysis using wavelet transform. *IAHS Publication*, v. 278, p. 195-201, 2003.

SANTOS, C. A. G.; IDEIÃO, S. M. A. Application of the wavelet transform for analysis of precipitation and runoff time series. *IAHS Publication*, v. 1, n. 303, p. 431-439, 2006.

SANTOS, C. A. G.; MORAIS, B. S.; SILVA, G.B.L. Drought forecast using Artificial Neural Network for three hydrological zones in San Francisco river basin. *IAHS Publication*, v. 333, p. 302-312, 2009.

SANTOS, C. A. G.; SRINIVASAN, V. S.; SUZUKI, K.; WATANABE, M. Application of an optimization technique to a physically based erosion model. *Hydrological Processes*, v. 17, n. 5, p. 989-1003, 2003. doi: 10.1002/hyp.1176.

SMITH, L. C.; TURCOTTE, D. L.; ISACKS, B. L. Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes*, v. 12, n. 2, p. 233-249, 1998. doi: 10.1002/(SICI)1099-1085(199802)12:2<233::AID-HYP573>3.0.CO;2-3.

TORRENCE, C.; COMPO, G. P. A practical guide to wavelet analysis. *Bulletin of the American Meteorological Society*, v. 79, n. 1, p. 61-78, 1998. doi: 10.1175/1520-0477(1998)079<0061:APGTWA>2.0.CO;2.

The Wavelet Transform And Its Application For Hydrological Time Series Analysis

ABSTRACT

This paper provides a practical guide of wavelet transform for hydrological data analysis. Thus, it is first applied with a hypothetical time series composed by four frequencies to better illustrate the technique, and then it is applied to two rainfall series from the rain gauges of São Francisco River basin for the period of 1936 to 2011 for the purpose of exploring this technique in the analysis of hydrological data. The wavelet transform corresponds to a tool that decomposes one-dimensional time series into a diffuse two-dimensional time-frequency image simultaneously, thereby determining the main modes of variability, and how they vary within the time. It is a recent advance in signal processing and is based on successive convolution of these functions with different scales, the wavelet functions. This paper is an explanation of the key features of wavelets and exemplifies some of their application, using rainfall data from São Francisco River basin, Brazil.

Keywords: Wavelet transform, signal processing, time series