ISSN:1984-2295

Revista Brasileira de Geografia Física

Homepage: <https://periodicos.ufpe.br/revistas/rbgfe>

Análise de Dados Climáticos Usando Transformada em Ondeletas Cruzada e Coerência

Roseilson Souza do Vale¹, Raoni Aquino de Santana¹, Cléo Quaresma Dias-Júnior²

¹ Dr. em Clima e Ambiente, Professor Adjunto, Instituto de Engenharia e Geociências, Universidade Federal do Oeste de Pará, Rua Vera Paz, s/n, CEP 68040-470, Santarém, Pará. (93) 99189-2970. roseilsondoval@gmail.com (autor correspondente). raoniass@gmail.com (segundo autor). ² Dr. em Clima e Ambiente, Professor Adjunto, Departamento de Física, Instituto Federal do Pará, Av. Almirante Barroso, s/n, CEP 66093-020, Belém, Pará. cleoquaresma@hotmail.com.

Artigo recebido em 27/08/2019 e aceito em 06/04/2020

RESUMO

Este estudo mostra uma análise em transformada em ondeleta cruzada e coerência aplicada a duas séries temporais, sendo uma delas precipitação e a outra temperatura do ar. O objetivo deste estudo é mostrar que esta técnica é uma ferramenta poderosa na análise de séries temporais climáticas, para isso aplicamos a duas séries com relação física muito conhecida na climatologia. Além da aplicação realizada, recorreu-se também a uma descrição matemática dos métodos. A técnica da transformada em ondeletas cruzada e coerência mostrou-se eficiente em capturar a relação matemática entre as séries de precipitação e temperatura do ar. Com este estudo esperamos difundir o uso desta técnica para fins de ensino e pesquisa em diversos sistemas geofísicos.

Palavras-chave: Ângulo de fase, Coerência em ondeleta, Ondelela cruzada, Precipitação, Temperatura.

Analysis of Climate Data Using Transformed Crosswave and Coherence

ABSTRACT

This study presents a cross wavelet transform and wavelet coherence analysis applied to a precipitation and an air temperature time series. The objective of this study is to demonstrate that this technique is a powerful tool in the analysis of climatic time series, and can be applied to two time series with very well-known physical relationships in terms of climatology. In addition to this application, a mathematical description of the methods was done. The cross-curves and coherence technique proved to be efficient in capturing the mathematical relationship between precipitation series and air temperature. With this study we hope to disseminate the use of this technique for teaching and research purposes in various geophysical systems.

Keywords: Phase Angle, Wavelet Coherence, Cross wavelet, Precipitation, Temperature.

Introdução

O uso da transformada em ondeletas (TO) para análise de séries temporais é uma ferramenta muito poderosa e amplamente difundida no meio científico. Sua aplicação é encontrada em diferentes áreas do conhecimento, tais como: hidrologia Labat et al. (2004) e Sang (2013), saúde (Addison et al., 2000; 2005), economia Conraria, Luís Aguiar, Azevedo; Joana (2008), ecologia Cazelles et al. (2008), turbulência atmosférica Dias-júnior et al. (2017 e Júnior (2008) e geofísica espacial Bolzan (2005). A difusão do uso da TO se deu, principalmente, após o trabalho de Farge

(1992) e do pacote criado por Torrence (1998). No entanto, o uso da TO cruzada aplicada a duas séries temporais, com a finalidade de analisar regiões comuns de potência, fase e correlação, tem seguido de modo mais lento. Essa ferramenta, nos permite ver muito além de uma simples correlação linear. Sua aplicação, assim como a TO aplicada a uma única série, pode ser feita de modo simples usando o pacote criado por Grinsted et al. (2004). Neste estudo aplicou-se a transformada em ondeleta cruzada e coerência em ondeleta a duas séries geofísicas, sendo elas: precipitação e temperatura

do ar. Recorreu-se a essas séries pois se conhece a relação física entre elas. Notou-se que o método utilizado foi robusto e eficiente na análise deste tipo de sinal. A descrição matemática da transformada em ondeleta contínua e transformada de Fourier, não foram abordadas aqui, uma vez que em Bolzan (2004, 2006) e Torrence (1998) já os fizeram de forma detalhada.

Material e métodos

Dados

Para exemplificar o método da transformada em ondeleta cruzada (TOC) e coerência em ondeleta (CO), foram usadas duas séries temporais geofísicas com efeitos conhecidos em nosso sistema climático: precipitação e temperatura do ar. Além de seus efeitos no tempo e no clima, a relação entre os dois sinais é bem estabelecida e consolidada. A fonte dos dados é do Banco de Dados Meteorológicos para Ensino e Pesquisa (BDMEP) do Instituto Nacional de Meteorologia (INMET), disponibilizado em <http://www.inmet.gov.br>. A série é composta por dados mensais de precipitação e temperatura do ar para a cidade de Manaus (AM) de 1961 à 2016.

Climatologia da área de estudo

Devido a sua localização próxima a linha do equador, a cidade de Manaus não possui as estações do ano bem definidas, como ocorre nas regiões sul e sudeste do Brasil. As estações são definidas basicamente em estação chuvosa (dez –

abr) e menos chuvosa (jul - out) Fisch, Gilberto; Marengo, José A.; Nobre (1998). A temperatura apresenta pouca variabilidade ao longo do ano, sendo mais marcante ao longo do dia. Assim, durante a estação chuvosa, a temperatura tende a diminuir, devido a presença de maior nebulosidade e consequentemente menos radiação chega à superfície. Por outro lado, durante a estação menos chuvosa a temperatura tende a ser maior comparada à estação chuvosa, devido a menor cobertura de nuvens e consequentemente mais radiação chegando à superfície. Outra variável que fortalece as elevadas temperaturas é a alta umidade relativa da região. A partir da relação física entre a precipitação e a temperatura do ar, fica claro que as respectivas séries temporais apresentam uma correlação negativa, ou seja, estão defasadas em 180° .

A Figura. 1 mostra a série temporal (parte superior) e o espectro de potência da TO para precipitação. De maneira análoga, a Figura.2 para a série temporal de temperatura do ar (a partir daqui apenas temperatura). Como esperado, as duas séries apresentam maior significância estatística na escala de um ano (região com contorno), devido ao ciclo anual bem definido, dos máximos e mínimos de precipitação e temperatura (Figura.1 e 2). O espectro de potência da precipitação apresenta uma descontinuidade entre 1978 – 1984 e a da temperatura entre 1978 – 1987 e 1990 – 1993. As regiões escuras dentro do cone de influência (COI) representam baixa significância no espaço tempo e frequência.

Figura 1. Série temporal de precipitação (superior) e seu espectro de potência em ondeleta (inferior) usando ondeleta de Morlet.

A linha de contorno mostra o nível de 5% de significância estatística contra o ruído vermelho.

Cone de influência (COI) onde os efeitos de borda podem distorcer a imagem, são mostrados com uma sombra mais clara. As regiões mais escuras (claras) representam menor (maior) significância

no espectro de potência, ou seja, menor (maior) variância para determinado período (escala) e tempo (ano).

Figura 2. Série temporal de temperatura (superior) e seu espectro de potência em ondeleta (inferior). A linha de contorno mostra o nível de 5% de significância estatística contra o ruído vermelho. Cone de influência (COI) onde os efeitos de borda pode distorcer a imagem, são mostrados com uma sombra mais clara. As regiões mais escuras (claras) representam menor (maior) significância no espectro de potência, ou seja, menor (maior) variância para determinado período (escala) e tempo (ano).

Como já mencionado, as séries apresentam uma correlação negativa, onde o coeficiente de correlação entre elas foi $r = -0.61$ e p-valor foi $p = 0.00$. Neste estudo, foi analisado o método da TOC usando para isso as séries temporais de precipitação e temperatura. Além disso, foi usado a CO para obtermos a relação de fase entre as séries em um escalograma de tempo – frequências fornecido pela TO. Caso o leitor não conheça ou queira se aprofundar no método da transformada em ondeleta, o mesmo pode fazer uma leitura fácil e agradável em Bolzan (2004, 2006) e Torrence (1998 e Torrence; Webster (1999), bem como nos sites <http://paos.colorado.edu/>, <http://www-stat.stanford.edu/> e <http://www.glaciology.net/>. As séries temporais foram analisadas usando a ondeleta de Morlet do pacote em Matlab criado por Grinsted et al. (2004).

Transformada em ondeleta cruzada

A TO é especialmente eficiente na análise de sinais transientes e não-estacionários, por permitir interpretação simultânea da informação temporal e espectral usando uma janela de comprimento variável. A TOC é eficiente para destacar regiões de energia coincidentes entre dois sinais no domínio da TO assim como determinar a fase relativa entre esses sinais. A TOC é

determinada pela multiplicação da primeira TO complexa com o complexo conjugado da segunda TO (Equação 1).

$$W_n^{XY}(s) = W_n^X(s)W_n^{Y*}(s) \quad (1)$$

Onde X e Y, representam as séries temporais, n é o tempo, s é a escala da ondeleta e (*) significa o complexo conjugado Torrence (1998 e Torrence; Webster (1999). A TOC é complexa, então podemos definir a espectro da ondeleta cruzada (EOC) como $|W_n^{XY}(s)|$. Enquanto, a TO revela que parte da série temporal analisada transporta energia significativa e em quais frequências isso acontece, a TOC exibe a covariância entre X e Y e revela informações sobre a relação entre as fases das séries. Isso permite que o EOC indique a quantidade de similaridade de potência entre X e Y.

A diferença de fase entre as séries X e Y é importante porque nos permite quantificar e qualificar a relação entre as séries em termos da propagação de seus ciclos de oscilação. Assim, é necessário estimar a média e o intervalo de confiança da diferença de fase entre as séries usando uma média circular (ou ângulo médio) de fase sobre as regiões com significância estatística acima de 5% que estão fora do COI para quantificar

a relação de fase Grinsted et al. (2004). A média circular do conjunto de ângulos (a_i , $i = 1 \dots n$) é definido como:

$$a_m = \tan^{-1}(X, Y) \quad (2)$$

$$X = \sum_{i=1}^n \cos(a_i) \quad (3)$$

$$Y = \sum_{i=1}^n \sin(a_i) \quad (4)$$

Devido os ângulos de fase não serem independentes, se torna difícil calcular o intervalo de confiança da média circular de forma confiável. Por isso, calculamos o espalhamento dos ângulos em torno da média, definido como desvio padrão circular:

$$s = \sqrt{-2 \ln(\sqrt{(X^2 + Y^2)}/n)} \quad (5)$$

A interpretação de s é análoga ao desvio padrão (σ), onde os ângulos estão distribuídos próximos da média circular. Em alguns casos, pode haver necessidade de se calcular o ângulo de fase médio para cada escala, então o ângulo de fase pode ser quantificado como um número de anos Grinsted et al. (2004).

Coerência em ondeleta

Como vimos no item 3.2, a TOC nos revela as regiões com significativa potência comum entre as séries. A coerência em ondeleta (CO) nos revela a coerência entre as TOC das duas séries em um diagrama tempo – frequência. Ou seja, explica quanto a variabilidade de uma determinada série é explicada por outra série. Segundo Torrence; Webster (1999), o quadrado da coerência em ondeleta é definido como o quadrado do valor absoluto da TOC suavizada, normalizada pelo espectro de potência em ondeleta,

$$R_n^2(s) = \frac{|(s^{-1}W_n^{XY}(s))|^2}{(s^{-1}|W_n^X(s)|^2)(s^{-1}|W_n^Y(s)|^2)} \quad (6)$$

onde $\langle * \rangle$ indica o operador suavização no tempo e em escala. O numerador possui a parte real e imaginária da TOC que são suavizados separadamente antes do cálculo do valor absoluto. O denominador possui o espectro de potência em ondeleta (depois de elevado ao quadrado) o qual será suavizado. O fator s^{-1} é usado para converter em densidade de energia. A partir destas definições, temos $R_n^2(s)$ variando entre 0 e 1.

Podemos observar que a Equação 6 é semelhante ao coeficiente de correlação de Pearson. Assim, podemos pensar na CO como sendo também a medida do coeficiente de

correlação localizado no tempo – frequência. O operador suavização é definido como:

$$\langle * \rangle = S_{escala}(S_{tempo}(W_n(s))) \quad (7)$$

onde S_{escala} é a suavização ao longo da escala dos eixos das ondeletas e S_{tempo} a suavização ao longo do tempo.

Para ondeleta de Morlet o operador de suavização é dado de acordo com Torrence (1998):

$$S_{tempo}(W)|_s = \left[W_n(s) * c_1^{-t^2/2s^2} \right]_s \quad (8)$$

$$S_{escala}(W)|_s = [W_n(s) * c_2 \Pi(0.6)]_s \quad (9)$$

onde c_1 e c_2 são as constantes de normalização e Π é a função retângulo. O fator 0.6 é determinado empiricamente para a ondeleta de Morlet Torrence (1998). Na prática as convoluções no tempo e em escala são feitas discretamente e portanto os coeficientes de normalização são determinados numericamente.

A medição da coerência de fase é realizada para atribuir um valor de suporte nas medidas de diferença de fase entre duas séries temporais, com grandes valores de coerência de fase, significando que a diferença de fase varia suavemente como uma função da frequência, ou seja, não é aleatório. A diferença de coerência de fase em ondeletas é dado segundo Torrence; Webster (1999):

$$\phi_n(s) = tg^{-1} \left(\frac{\Im\{(s^{-1}W_n^{XY}(s))\}}{\Re\{(s^{-1}W_n^{XY}(s))\}} \right) \quad (10)$$

As partes real ($\Re$) e imaginária ($\Im$) suavizada já poderiam ser calculadas na Equação 9. Tanto $R_n^2(s)$ quanto $\phi_n(s)$ são funções do índice de tempo n e de escala s . A suavização na Equação 9 e Equação 10 é realizada usando convolução na direção do tempo e escala.

No pacote criado por Grinsted et al. (2004) e usado neste estudo, o nível de significância estatística da CO é estimada usando o método de Monte Carlo. É gerado um grande conjunto de pares de conjuntos de dados de substituição com os mesmos coeficientes AR1 (modelo autoregressivo) que os conjuntos de dados de entrada. Para cada par é calculado a CO. Então, é estimado o nível de significância para cada escala usando apenas valores fora do COI. Para compensar o pouco impacto dos coeficientes de AR1 no nível de significância, a estimativa de Monte Carlo é da ordem de 1000 pares de conjuntos de dados de

substituição. O número de escalas por octave é suficientemente alto para capturar a forma retangular do operador de suavização de escala enquanto o tempo de computação é minimizado.

Resultados e discussão

O EOC associada as duas TO de precipitação e temperatura descritas nas Figura. 1 e 2 são mostradas na Figura 3. Notamos que as características comuns vistas nas TO individuais de precipitação e temperatura (Figura. 1 e 2) destacam-se como sendo significante estatisticamente. Com apenas uma falha na significância entre 1979 – 1984. Contudo, percebe-se que as séries encontram-se defasadas em 180° (as setas apontam para a esquerda). Uma

Empiricamente, o pacote de Grinsted et al. (2004) encontra 10 escalas por octave para ser satisfatório.

defasagem praticamente constante em todas as escalas sugere um atraso temporal constante devido ao mecanismo físico da precipitação e temperatura em Manaus. O que era de se esperar e já fora justificado anteriormente. Assim, é satisfatório que a TOC mostre que a precipitação e a temperatura estejam em anti-fase em todos os anos com significante espectro de potência comum para a escala de um ano. No entanto, mesmo para as demais escalas, fora da significância estatística (regiões claras), a TOC também mostra uma relação predominante de anti-fase.

Figura 3. Espectro da ondeleta cruzada entre as séries temporais de precipitação e temperatura. A linha de contorno mostra o nível de 5% de significância estatística contra o ruído vermelho. A relação entre as fases relativas das séries são mostradas como setas que, apontando para a direita indicam fase (0°), e apontando para a esquerda indicam anti-fase (180°). As regiões mais escuras (claras) representam menor (maior) significância no espectro de potência, ou seja, menor (maior) variância para determinado período (escala) e tempo (ano).

Uma defasagem de 180° nos diz que o pico máximo da temperatura ocorre 6 meses após o pico máximo da precipitação, ou que o quando uma série alcança seu máximo a outra está no seu mínimo. A TOC entre as séries está de acordo com climatologia de Manaus, onde o mês mais chuvoso ocorre em março e a temperatura máxima em setembro, o que nos dá uma defasagem de 6 meses. O desvio de quase total anti-fase sugere que a precipitação conduz a temperatura. Outra informação interessante que a TOC fornece é a relação de fase entre as séries em diferentes escalas

e ao longo dos anos, diferente de uma correlação linear, que fornece apenas um valor. Ou seja, duas séries podem apresentar um fraco coeficiente de correlação mas quando decompõe-se as séries em escalas usando a TOC, é possível encontrarmos forte correlação entre elas Addison (2004).

A coerência em ondeleta elevada ao quadrado ($R_n^2(s)$) entre as séries de precipitação e temperatura é mostrado na Figura. 4. Percebemos que comparado a TOC (Figura. 3) a CO apresenta uma ampla região com significância estatística (linha de contorno) e não apenas para a escala de

um ano, como na Figura 3. Praticamente, todas as regiões em destaque apresentam uma relação de anti-fase entre as séries de precipitação e temperatura. A Figura 4 apresenta ainda uma

É possível que duas séries estejam perfeitamente correlacionadas em uma escala específica, enquanto a área de correlação significativa é muito inferior a 5%. No entanto, a região significativa fora do COI é relativamente extensa, sendo pouco provável que isso seja simplesmente por acaso. Já era de se esperar que a

região de destaque entre 1961 e 1984, no entanto, está fora do COI. A área de um gráfico frequência - tempo acima do nível de significância de 5% não é uma indicação confiável de causalidade.

correlação entre as séries de precipitação e temperatura fosse elevada na escala de um ano, como já discutido no item 3.2. No entanto, é possível observar que existe uma outra região na escala de 2 a 6 anos com forte coerência entre os anos de 1971 a 2016. Ou seja, além da relação entre as séries na escala sazonal, existe uma forte correlação na escala interanual.

Figura 4. Coerência em ondeleta elevada ao quadrado ($R_n^2(s)$) entre as séries de precipitação e a temperatura para a cidade de Manaus. A linha de contorno mostra o nível de 5% de significância estatística contra o ruído vermelho. A relação entre as fases relativas das séries são mostradas como setas, onde setas apontando para a direita – as séries estão em fase (0°), setas apontando para a esquerda – as séries estão em anti-fase (180°). As regiões mais escuras (claras) representam menor (maior) significância no espectro de potência, ou seja, menor (maior) variância para determinado período (escala) e tempo (ano).

Conclusões

Foram analisadas séries temporais de precipitação e temperatura via transformada em ondeletas contínua de Morlet. A transformada em ondeletas cruzada nos revelou regiões comuns com significativa potência e a relação de fase entre as séries. Já a coerência em ondeletas apresenta a capacidade de mostrar a correlação local entre duas transformada em ondeletas cruzada em tempo e frequência. O propósito deste estudo foi mostrar a aplicação deste método muito poderoso na análise de sinais geofísicos. Para o exemplo usado, o método mostrou-se robusto e eficiente, capturando a relação matemática das séries de precipitação e temperatura, e justificado pela relação física entre

as mesmas. O uso desta ferramenta torna-se prático e fácil, tanto para o ensino quanto pesquisa, uma vez que encontra-se disponibilizado livremente na web para seu uso. Com isto, espera-se auxiliar e incentivar o estudo de séries temporais, usando a transformada em ondeletas cruzada e coerência em ondeletas, aplicado a ampla diversidade de sinais geofísicos que temos.

Referências

- Addison, P. S. 2004. The illustrated wavelet transform handbook.
- Addison, P. S. 2005. Wavelet transforms and the ECG: a review. *Physiological Measurement*, 26, R155–R199.

- Addison, P. S.; Watson, J. N.; Clegg, G. R.; et al. 2000. Evaluating Arrhythmias in ECG Signals Using Wavelet Transforms. *IEEE Engineering in Medicine and Biology*, September/October, p. 0–5.
- Bolzan, M. J. A. 2004. Análise da transformada em ondeletas aplicada em sinal geofísico. *Revista Brasileira de Ensino de Física*, 26, 37–41.
- Bolzan, M. J. A. 2005. Statistical and Wavelet Analysis of the Solar Wind Data. *Brazilian Journal of Physics*, 35, 592–596.
- Bolzan, M. J. A. 2006. Transformada em ondeleta: uma necessidade. *Revista Brasileira de Ensino de Física*, 28, 563–567.
- Cazelles, B.; Chavez, Æ. M.; Berteaux, Æ. D.; et al. 2008. Wavelet analysis of ecological time series. *Oecologia*, p. 287–304.
- Conraria, L. A., Azevedo, N.; Joana, M. 2008. Using wavelets to decompose the time – frequency effects of monetary policy. *Physica A: Statistical Mechanics and its Applications*, 387, 2863–2878.
- Dias-júnior, C. Q. 2008. Estudo Comparativo da Estrutura da Turbulência Atmosférica na Camada Limite Superficial Acima da Floresta Primária e Região Desmatada da Amazônia : análise de ondeletas aplicada à detecção de estruturas. p. 104.
- Dias-júnior, C. Q.; Sá, L. D. A.; Marques, E. P.; et al. 2017. Agricultural and Forest Meteorology Turbulence regimes in the stable boundary layer above and within the Amazon forest. *Agricultural and Forest Meteorology*, 233, 122–132.
- Farge, M. 1992. Wavelet Transforms and their Applications to Turbulence. *Annual Review of Fluid Mechanics*, 24, n. 1, p. 395–458.
- Fisch, G.; Marengo, J. A.; Nobre, C. A. 1998. Uma Revisão geral sobre o Clima da Amazônia. *Acta Amazonica*, 28, 101–126.
- Grinsted, A.; Moore, J. C.; Jevrejeva, S. 2004. Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, 11, 561–566.
- Labat, D.; Ronchail, J.; Callede, J.; et al. 2004. Wavelet analysis of Amazon hydrological regime variability, 31, 3–6.
- Sang, Y. 2013. A review on the applications of wavelet transform in hydrology time series analysis. *Atmospheric Research*, 122, 8–15.
- Torrence, C. and G. P. C. 1998. A practical guide to wavelet analysis. *Bull Am Meterol Soc*, 79, 61–78.
- Torrence, C.; Webster, P. J. 1999. Interdecadal Changes in the ENSO – Monsoon System. *Journal of Climate*, 12, 2679–2690.