



CORRELAÇÃO MÚLTIPLA WAVELET

Vinícius Basseto Félix (PIBIC/CNPq/FA/Uem), Eniuce Menezes de Souza (Orientador), e-mail: felix_prot@hotmail.com.

Universidade Estadual de Maringá / Centro de Ciências Exatas, PR.

Ciências Exatas e da Terra/ Probabilidade e Estatística.

Palavras-chave: *Wavelet*, Séries Temporais, Correlação Múltipla.

Resumo:

Neste projeto foi estudada a correlação múltipla *wavelet*, do inglês, *wavelet multiple correlation* (WMC), e sua aplicação em séries temporais (ST). A WMC proporciona vantagens significativas em relação à correlação cruzada amostral, principalmente se aplicada em mais de duas séries, pois não só facilita e condensa as informações para análise, como também permite a visualização de comportamentos que antes não poderiam ser identificados. Além disso é capaz de lidar com processos não-estacionários e permite uma análise multiescalar.

Introdução

Este estudo é uma continuação do projeto de iniciação científica “Tópicos em Variância *Wavelet*”, no qual estudou-se as *wavelets*, desde seus princípios, até a análise de variância construída a partir de *wavelets* para análise univariada de ST. Iniciou-se a investigação para análise bivariada de ST a partir da correlação cruzada *wavelet* (WCC - *wavelet cross correlation*). Além da implementação da teoria estudada também foram feitas aplicações em dados simulados e reais. Busca-se agora neste projeto a expansão da WCC para o caso multivariado (WMC), a fim de checar e comparar a metodologia com as mais usuais, comprovar suas vantagens, justificar e incentivar seu uso.

Basicamente, a WMC analisa de forma multiescalar a ST no domínio tempo-escala, assim sua aplicação é facilmente feita em diversas áreas. Trata-se





de um desenvolvimento recente, mas que já tem propiciado resultados importantes.

Materiais e métodos

Como visto em Tiwari et al. (2013), tomando dois processos estocásticos (X_{1t} e X_{2t}), sabe-se que para o desenvolvimento da correlação *wavelet* é necessária a construção das variâncias de $\{X_{1t}\}$ e $\{X_{2t}\}$ em diferentes escalas e da covariância $\{X_{1t}, X_{2t}\}$. A variância *wavelet* então é estimada a partir da *maximal overlap discrete wavelet transform* (MODWT) e seus coeficientes para escala $\lambda_j = 2^{j-1}$, sendo expressa por

$$\hat{\sigma}_x^2 = \frac{1}{\hat{n}_j} \sum_{k=L_j-1}^{n-1} (\hat{W}_{j,k})^2, \quad (1)$$

em que

- $\hat{W}_{j,k}$ é o coeficiente da MODWT da variável X na escala λ_j ;
- $\hat{n}_j = L_j + 1$ é o número de coeficientes não afetados pela fronteira;
- $L_j = (2^j - 1)(L - 1) + 1$ é o tamanho do filtro *wavelet* na escala λ_j .

A covariância *wavelet* decompõe a covariância entre dois processos estocásticos entre escalas, e pode ser escrita como:

$$Y_{x_1 x_2} = cov_{x_1 x_2}(\lambda_j) = \frac{1}{\hat{n}_j} \sum_{k=L_j-1}^{n-1} (\hat{W}_{j,k})^{x_1} (\hat{W}_{j,k})^{x_2}. \quad (2)$$

Dado (1) e (2) pôde-se expressar o estimador MODWT da correlação *wavelet* dado por

$$\hat{\rho}_{x_1 x_2} = \frac{cov_{x_1 x_2}(\lambda_j)}{\hat{\sigma}_{x_1}^2 \hat{\sigma}_{x_2}^2}, \quad (3)$$

Porém, esta comparação par a par possui diversas limitações, e assim a WMC mostra-se extremamente útil. Considerando agora $X_t = (x_{1t}, x_{2t}, \dots, x_{mt})$ um processo estocástico multivariado e $W_{jt} = (w_{1jt}, \dots, w_{mjt})$ os respectivos coeficientes *wavelet* numa escala λ_j obtidos através da aplicação da MODWT para cada processo x_{it} .

A WMC, ou $\varphi_x(\lambda_j)$, pode ser definida como

$$\varphi_x(\lambda_j) = \sqrt{1 - \frac{1}{\max \text{diag } F_j^{-2}}}. \quad (4)$$





em que

- P_j é a matriz de correlação de $W_{j,t}$;
- $\max \text{diag}(\cdot)$ é o operador que seleciona o maior elemento da diagonal do argumento.

O intervalo de confiança é dado por

$$\tilde{\varphi}_x(\lambda) \pm Z_{\frac{\alpha}{2}} \left[\frac{\hat{\sigma}_x}{n_j} \right]^2.$$

Resultados e Discussão

Para aplicação em dados reais foram usadas séries temporais do índice de cintilação S4 no período de fraco (Março, 07/04/2013) e forte (Junho, 07/04/2013) cintilação ionosférica, que ocorre quando o sinal passa por irregularidades na ionosfera. O efeito de cintilação ionosférica é o principal obstáculo atual para que aplicações de alta acurácia envolvendo posicionamento por satélite como os sinais dos satélites do *Global Positioning System* (GPS) sejam plenamente operacionais, tais como aterrissagem de aviões e agricultura de precisão, por exemplo.

Assim, para cada análise (período fraco e forte), sete séries temporais diárias foram observadas com intervalo de um minuto para o satélite 11 na estação PRU1, Presidente Prudente, São Paulo. Os dados foram obtidos do banco de dados CIGALA (<http://iscigalacalibra.fct.unesp.br/is/ismrtool/mining/Mining.php>).

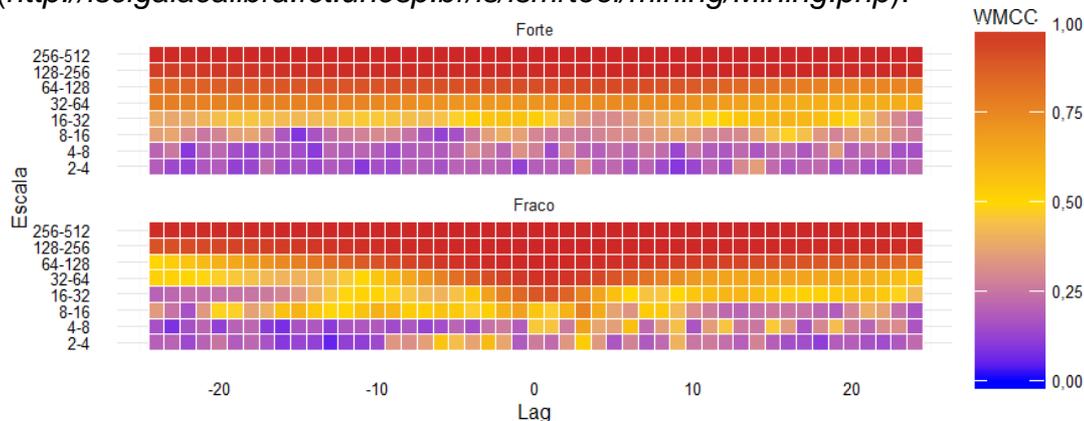


Figura 1 – WMC das sete ST para o índice S4, dos períodos forte e fraco de cintilação fraca, para cada escala e diferentes lags.





Pela Figura 1, vê-se que correlações fortes aparecem nas maiores escalas que incluem os efeitos ocorrendo num período de 346 minutos que representa o comportamento diário em forma de U também correlações no $lag 0$ são, geralmente, mais fortes.

Como a cintilação é muito fraca neste período, outros efeitos sistemáticos, como o efeito de multicaminho devido a reflexões de sinais em superfícies próximas à antena, são dominantes no índice S4, como discutido em Brassarote (2014). Nota-se também na Figura 1 que as maiores correlações aparecem na escala de período 8-16 e em lags entre 3 e 4 minutos devido a diferença de 236 segundos entre dias sideral e solar, o que faz com que o efeito do multicaminho se repita diariamente, mas com tal defasagem.

Este resultado corrobora com as suspeitas levantadas em Brassarote (2014) sobre a presença de efeitos como o multicaminho que se repete a cada dia se o ambiente receptor permanecer inalterado. Além disto, após tal identificação, este efeito poderá ser corrigido nos índices S4 para que os mesmos sejam mais realísticos na representação da cintilação ionosférica.

Conclusões

Portanto, fica evidente as vantagens de trabalhar-se em multiescala, a qual permite a análise completa de séries temporais, que ficaria muitas vezes mal interpretada, caso fossem usadas metodologias tradicionais, por sua incapacidade de capturar todos os efeitos nas ST.

A WMC, permite a visualização geral da correlação em diferentes escalas e $lags$ através de um único gráfico. Isto foi evidenciado nos *heatmaps* (Figura 1). O que torna comportamentos e efeitos de difícil identificação agora visíveis, como visto na aplicação, em que se nota o efeito de multicaminho.

Agradecimentos

Agradecemos ao CNPq-FA-UEM, por ter possibilitado e financiado este projeto.

Referências

Brassarote, G. d. O. N. (2014). **Análise multiescala de séries temporais do efeito da cintilação ionosférica nos sinais de satélite GPS a partir de wavelets não decimadas.**

Tiwari, A. K., Andries, M. A., Ihnatov, I. (2013). **Analysis of CEEC exchange rates comovements using wavelet multiple correlation and cross-correlation.** Available at SSRN 2276380.

