A multiresolution and multivariate analysis of the dynamic relationships among energy time series

Josué M. Polanco-Martínez[1,2]*, Luis M. Abadie[1], Javier Fernández-Macho[2]

[1]University of the Basque Country & [2]Basque Centre for Climate Change

*E-mail: *josue.m.polanco@gmail.com

Abstract

This work proposes the use of a novel multivariate, dynamic approach – wavelet local multiple correlation (WLMC) (Fernández-Macho, 2018) – to analyse the relationship between energy time series in the time-scale domain. This approach is suitable for use with any kind of data that change over time and involve heterogeneous agents who make decisions across different time horizons and operate on different time scales. To exemplify the use of WLMC, we analyse the relationships between the prices of seven commodities: West Texas Intermediate (WTI) crude oil and six distilled products (conventional gasoline, regular gasoline, heating oil, diesel fuel, kerosene and propane) from 2006 to 2017. The results reveal that the wavelet correlations are strong throughout the period studied and there is a strong decay in correlation values from 2013 to 2015. The most plausible explanation for this decay is overproduction of tight oil in the U.S. and a slowdown in global demand for oil. Furthermore, our results also reveal that heating oil, diesel and kerosene maximise the multiple correlation with respect to the other oil variables on different scales, indicating that these products are the most dependent variables in the crude-product/price system. WLMC offers new opportunities for applications in energy research and other fields such as climatology, geophysics, etc.

Introduction and motivation

The study of the relationships between the prices of crude oil and refinery products is of great interest to refiners, investors, policymakers and scholars. One of the main reasons for that interest is that it is extremely important to know the differences or margins between the output prices of distilled products for petroleum industry refiners and crude oil costs. It is therefore fundamental to investigate what dynamic interactions there are between crude and refined product prices. Despite the importance of the study of the links between the prices of crude oil and its products, the number of papers published from 2005 to the beginning of 2010 is small, especially as regards papers in which several products are analysed, in comparison, for example, to the number of studies published about the link between crude oil spot prices and futures markets [8].

During the last few years, however, there has been a relative increase in the number of publications that threat the prices of crude oil and the main products as a system, and in particular in paper that use a variety of advanced statistical techniques originally developed to analyse complex systems. For instance, [14] uses a copula approach to capture potential nonlinear relations between crude oil and some refined products. [7] uses detrended cross-correlation analysis to investigate the cross-correlations between crude oil and refined product prices for 1991–2013. [17] also investigates the same problem and finds that nonlinear correlations are stronger in the long-term than in the short-term, crude oil and product prices are co-integrated and financial crisis in 2007–2008 caused a structural break in the co-integrating relationship. [2] analyses a dynamic conditional correlation between crude oil and fuels prices in a non-linear framework.

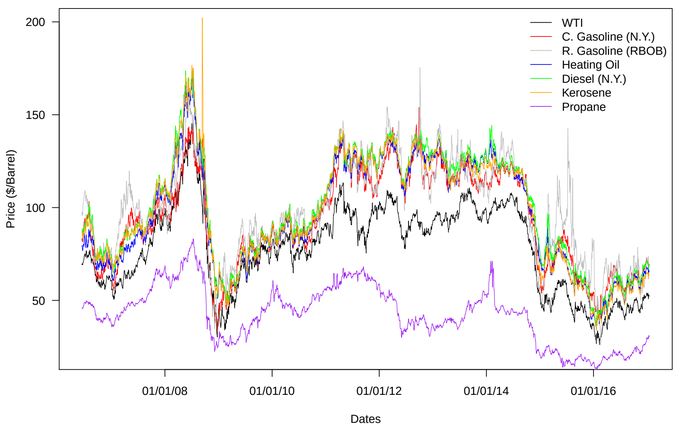
Such statistical techniques are suitable way of obtaining knowledge in energy research due the fact that oil/petroleum products prices make up a complex system [16, 18]. This means that there are a large number of variables and factors interacting with one another and belonging to a range of areas that are interconnected, e.g. energy industry, economics, finance, engineering, technology, environment, etc. [1, 9, 12]. On the other hand, it is necessary that the statistical techniques used to analyse energy time series are able to tackle many of the characteristics contained in energy data, such as: 1) energy time series change over time, i.e., they are not stationary (their mean or variance can change over time); 2) these series are not necessarily normally distributed [3, 17]; 3) they can display nonlinear structures [11, 17]; and 4) and they can involve heterogeneous agents who make decisions with different time horizons and operate on different time scales [11, 13]. An suitable tool to address many of these features is the wavelet local multiple correlation (WLMC) [4].

Objective

The objective of this work is to use a novel multivariate, dynamic approach – wavelet local multiple correlation (WLMC) [4] – to analyse the link between energy time series in the time-scale domain. As a case study, we analyse dynamically the links between the prices of seven commodities: crude oil prices (West Texas Intermediate spot prices, one of the world's main crude oil benchmarks) and the prices of six refined products (conventional gasoline, RBOB regular gasoline, heating oil, Ultra-Low-Sulphur diesel fuel, kerosene and propane) using daily market prices from 14/06/2006 to 17/01/2017. We are specially interested in the role of the subprime financial crisis and the growing tight oil production.

Materials and Methods

The data used in this study comprise daily market quotes for West Texas Intermediate (WTI), conventional gasoline (N.Y.), RBOB regular gasoline, heating oil, Ultra-Low-Sulphur diesel Fuel (N.Y.), kerosene and propane. All data sets cover the period from 16/06/2006 to 17/01/2017 (2668 trading days). Data were obtained from the U.S. Energy Information Administration (EIA).



Mathematical Section

Since each $x_i(t)$ is a discrete time series of length T for most applications, it can be shown that each $T \times n$ matrix of wavelet coefficients $W(j)$ can be obtained from

$$W(j) = \mathcal{W}_j X, \quad j = 1 \dots J, \tag{1}$$

where $X = [x(1), \cdots, x(T)]^\top$, $J = \log_2(T)$ is the maximum scale level and the rows of matrix $\mathcal{W}_j$ are made up of circularly shifted versions of the vector of T wavelet filter values $\psi_{j\tau}$ from the CWT (continuous wavelet transform) for a given discrete wavelet filter such as any member of the Daubechies family [10]. It can also be noted that at each level j these wavelet coefficients are associated with changes in the effective scale of length $\lambda_j = 2^{j-1}$, which roughly corresponds to periods in the range of $[2^j, 2^{j+1}]$ time units, while the "smooth" remainder is contained in one scaling coefficient [15]. Apart from its time redundancy, MODWT is known to have several major advantages, including energy preservation, which is particularly important in many applications ([10], [6, p.135]). The MODWT wavelet coefficients thus obtained from the basis for the wavelet multiple regression statistics calculated by the WLMC.

Consider $W(j) = [w_1(j), w_2(j), \dots, w_n(j)]$ in (1), i.e. the matrix of scale λ_j wavelet coefficients obtained by applying MODWT to each series in a realization of a multivariate stochastic process, and also define, $W_i(j)$ at each scale λ_j as the $T \times n$ matrix obtained by replacing the transform w_i in $W(j)$ by a column of ones. The idea is to obtain a linear function at each scale λ_j that minimises a weighted sum of squared errors for a fixed $\tau \in [1, \dots, T]$

$$S_{j\tau} = u_{j\tau}^\top \Theta_\tau u_{j\tau} \quad \text{with} \quad u_{j\tau} = W_i(j)\,\beta_{j\tau} - w_i(j), \quad \forall j, \tag{2}$$

where Θ_τ is the $T \times T$ diagonal matrix whose diagonal elements are the values $\theta(t-\tau)$ for a given moving average weight function or window $\theta(d)$ that depends on the time lag between observed MOWDT coefficients $w_i(j,t)$ and $w_i(j,\tau)$.

Letting τ move over time, T local linear regression fits are obtained at each scale λ_j, each with its corresponding residual weighted sum of squares

$$\hat{S}_{j\tau} = \hat{u}_{j\tau}^\top \Theta_\tau \hat{u}_{j\tau}, \quad \hat{u}_{j\tau} = W_i(j)\,\hat{\beta}_{j\tau} - w_i(j), \quad \forall j, \tau. \tag{3}$$

where $\hat{\beta}_{j\tau}$ are the least-squares estimates of the unknown parameters $\beta_{j\tau}$ in (2). These residuals can then be used to calculate J series of local coefficients of determination

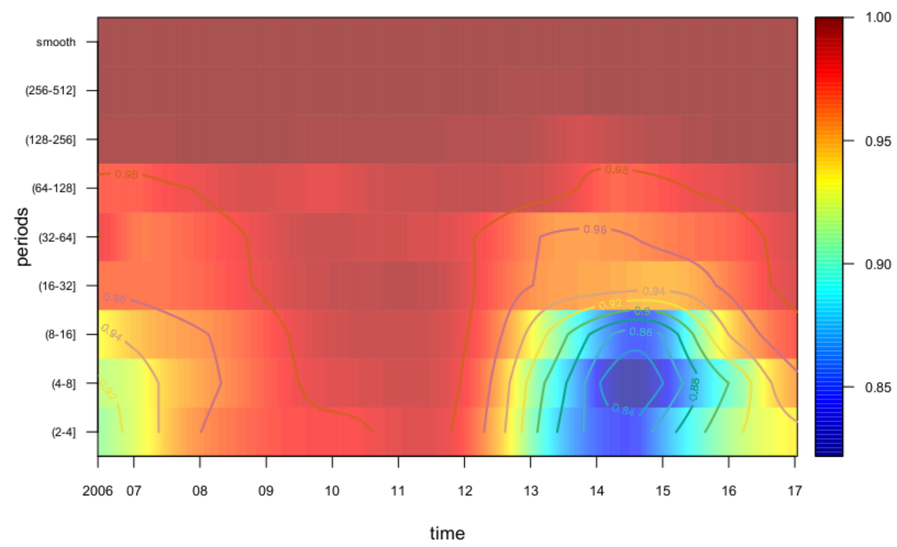
$$R^2_{j\tau} = 1 - \hat{S}_{j\tau}/TS_{j\tau}, \quad \tau = 1 \dots T,\ j = 1 \dots J, \tag{4}$$

where $TS_{j\tau}$ is the total weighted sum of squares at time τ and scale λ_j that measure how good the least-squares fit in (2) at each time τ and scale λ_j.

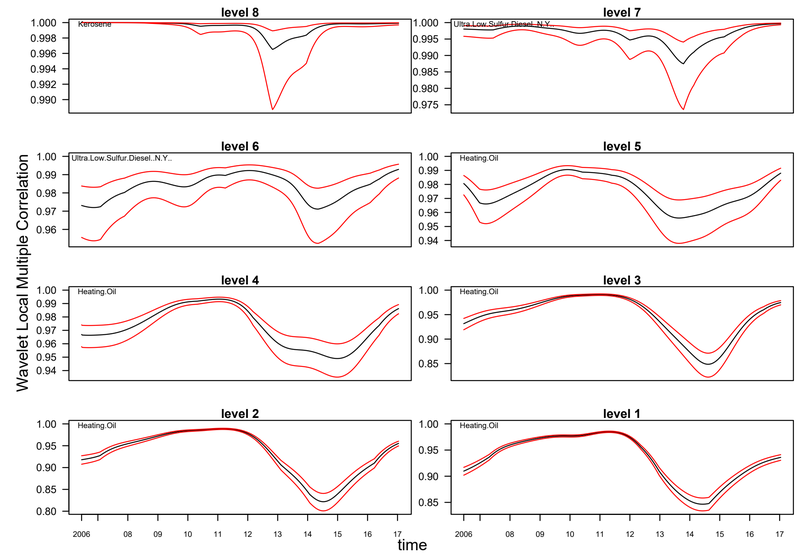
Finally, the WLMC is obtained as a collection of local multiple correlation coefficients $\varphi_X(\lambda_j, \tau)$ calculated as the square roots of each of the coefficients of determination $R^2_{j\tau}$ corresponding to the linear combination of weighted variables $\Theta_\tau^{1/2} w_{ij}$, i=1…n, where that local coefficient of determination is a maximum.

To estimate the WLMC we use the *wavemulcor* R package, which is freely available from the CRAN repository [5].

Results



1. Wavelet correlations are strong practically throughout the period under study (the correlation coefficients have values ranging from 0.82 to 0.99) for all the scales increasing from shorter to longer scales, particularly the last two wavelet scales (values of almost 1). These scales are associated with time horizons of between 64 and 128 days and scales (periods) between twice- and two-yearly.
2. The most conspicuous result is the strong decay of the wavelet correlation values that takes place approx. from 2013 to 2015 in the first three wavelet scales. There are two potential explanations:
   (a) Since 2014 WTI crude oil prices have been influenced enormously by non-conventional crude oil production in the U.S.
   (b) This abrupt decline could been driven by a combination of positive oil supply shocks, negative shocks to the storage demand for oil and negative shocks to consumption demand associated with an unexpected slowing of the global economy.
3. Our results also reveals that for the first five scales heating oil maximises the multiple correlation of the other variables, i.e., heating oil can be determined as a linear combination of the rest of the variables for these scales, indicating that heating oil is the most dependent variable in the crude/product price system.



Conclusions

- This work **proposes the use of an innovative multivariate and dynamic approach** – wavelet local multiple correlation (WLMC) [4] – **to analyse the relationship between energy time series in the time-scale domain**. This approach is suitable for use with any kind of data that change over time (non-stationary) and involve heterogeneous agents who make decisions on different time horizons and operate on different time scales.
- The WLMC results reveal that the bf correlations are strong, for practically the whole period under study, and **particularly for the two longest scales**. However, the most interesting result is the **strong decay in wavelet correlation values that takes place approx. from 2013 to 2015** for the first three wavelet scales. The most plausible explanation for this decay is overproduction of non-conventional crude oil in the U.S. and a slowdown in global demand for oil – probably caused by the recent global financial crisis.
- The WLMC also is able to provide a very interesting and original feature, i.e, it is **highlighted the variable that maximises the multiple wavelet local correlation in each wavelet scale**. Following this characteristic of the WLMC, our WLMC results also reveal that heating oil maximises multiple correlation with other oil variables for the first five scales, whereas diesel and kerosene are the variables that maximises multiple correlation with the rest of variables at the longest scales. This means that the **heating oil, diesel and kerosene are the most dependent variable in the crude/product prices system**.

This work is based on our recently published paper: Polanco-Martínez, J. M.*, Abadie, L. M., Fernández-Macho, J. (2018). A multi-resolution and multivariate analysis of the dynamic relationships between crude oil and petroleum-product prices. *Applied Energy*, 228, 1550-1560.

References

[1] J. Alvarez-Ramirez, M. Cisneros, C. Ibarra-Valdez, and A. Soriano. Multifractal Hurst analysis of crude oil prices. *Physica A: Statistical Mech. and its Applic.*, 313(3):651–670, 2002.
[2] A. Block S., M. B. Righi, S. G. Schlender, and D. A. Coronel. Investigating dynamic conditional correlation between crude oil and fuels in non-linear framework: The financial and economic role of structural breaks. *Energy Economics*, 49:23–32, 2015.
[3] R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1(2):223–236, 2001.
[4] J. Fernández-Macho. Time-localized wavelet multiple regression and correlation. *Physica A: Statistical Mechanics and its Applications*, 492:1226–1238, 2018.
[5] Javier Fernndez-Macho. wavemulcor: Wavelet routines for global and local multiple correlation, 2017. R package version 2.1.0.
[6] R. Genay, F. Seluk, and B. Whitcher. *An introduction to wavelets and other filtering methods in finance and economics*. Academic Press, 2002.
[7] L. Liu and G. Ma. Cross-correlation between crude oil and refined product prices. *Physica A: Statistical Mechanics and its Applications*, 413:284–293, 2014.
[8] A. Murat and E. Tokat. Forecasting oil price movements with crack spread futures. *Energy Economics*, 31(1):85–90, 2009.
[9] E. Panas and V. Ninni. Are oil markets chaotic? A non-linear dynamic analysis. *Energy Economics*, 22(5):549–568, 2000.
[10] D.B. Percival and A.T. Walden. *Wavelet methods for time series analysis*. Cambridge University Press, 2000. reprinted 2006.
[11] J. M. Polanco-Martínez and L.M. Abadie. Analyzing crude oil spot price dynamics versus long term future prices: A wavelet analysis approach. *Energies*, 9(12):1089, 2016.
[12] A Suleymanov, A Abbasov, and A Ismaylov. Fractal analysis of time series in oil and gas production. *Chaos, Solitons & Fractals*, 41(5):2474–2483, 2009.
[13] A.K. Tiwari and C.T. Albulescu. Oil price and exchange rate in india: Fresh evidence from continuous wavelet approach and asymmetric, multi-horizon granger-causality tests. *Applied Energy*, 179:272–283, 2016.
[14] B. Tong, C. Wu, and C. Zhou. Modeling the co-movements between crude oil and refined petroleum markets. *Energy Economics*, 40:882–897, 2013.
[15] B. Whitcher, P. Guttorp, and D.B. Percival. Wavelet analysis of covariance with application to atmospheric time series. *Journal of Geophysical Research*, 105:941–962, 2000.
[16] T. Yao and C-Q Zhang, Y-Jand Ma. How does investor attention affect international crude oil prices? *Applied Energy*, 205:336–344, 2017.
[17] T. Zhang, G. Ma, and G. Liu. Nonlinear joint dynamics between prices of crude oil and refined products. *Physica A: Statistical Mechanics and its Applications*, 419:444–456, 2015.
[18] Y-J Zhang. Speculative trading and WTI crude oil futures price movement: an empirical analysis. *Applied Energy*, 107:394–402, 2013.