Food, Fuels or Finances: Which Prices Matter for Biofuels (Supported by EU 7FP Project ECOCEP)

Ondrej Filip[2] Karel Janda[1,2] Ladislav Kristoufek[2,3]
David Zilberman[4]

[1]University of Economics, Prague
[2]Charles University in Prague
[3]Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic
[4]University of California, Berkeley

November 2016

Outline

Our contribution

- Focus on price relationships
- Graph theory methodology for finding price connections
- Wavelet coherence analysis of biofuels and related commodities

Outline

Previous research

- Review in Serra and Zilberman (2013 EE), book by de Gorter, Drabik, Just (2015)
- Cointegration studies (Serra et al., 2011, energy prices and agr. commodity prices correlated)
- VAR studies (McPhail, 2011: higher demand for ethanol $\rightarrow$ lower crude oil prices)
- mGARCH studies (Zhang, 2009, 2010: no long run relationship among prices of ethanol, corn, gasoline)
- Causality and predistability in distribution (Bastianin, Galeotti, and Manera, 2014: causality from crops to ethanol, not vice versa)

Outline

Motivation

- Use a simple methodology and describe the basic interactions between biofuels, their production factors (feedstock) and related fossil fuels.
- Use a model free approach (results seem to be quite model dependent)

$\rightarrow$ Use minimal spanning trees and hierarchical trees to uncover the most important and stable connections in the biofuels network (practically no assumptions with the exception of stationarity).
$\rightarrow$ Follow up with wavelets.

Outline

Correlations

- For a pair of assets i and j with values X_i and X_j and $i, j = 1, \ldots, T$, the sample correlation coefficient $\widehat{\rho}_{ij}$ is calculated as

$$\widehat{\rho}_{ij} = \frac{\sum_{t=1}^{T} (X_{it} - \overline{X_i})(X_{jt} - \overline{X_j})}{\sqrt{\sum_{t=1}^{T} (X_{it} - \overline{X_i})^2 \sum_{t=1}^{T} (X_{jt} - \overline{X_j})^2}}, \tag{1}$$

  where $\overline{X_i} = \frac{\sum_{t=1}^{T} X_{it}}{T}$ and $\overline{X_j} = \frac{\sum_{t=1}^{T} X_{jt}}{T}$ are respective time series averages.
- For a portfolio of N assets, we obtain $N(N-1)/2$ pairs of correlations.

Distance measure

- Mantegna (1999) showed that the correlation coefficients can be transformed into distance measures, which can in turn be used to describe hierarchical organization of the group of analyzed assets. Distance measure

$$d_{ij} = \sqrt{2(1 - \rho_{ij})} \quad (2)$$

  fulfills three axioms of a metric distance:
  - $d_{ij} = 0$ if and only if $i = j$;
  - $d_{ij} = d_{ji}$;
  - $d_{ij} \leq d_{ik} + d_{kj}$ for all k
- From the definition of the correlation coefficient, the distance ranges between 0 and 2, while $d_{ij} \to 0$ means that the pair is strongly correlated, $d_{ij} \to 2$ implies strongly anti-correlated pair and $d_{ij} = \sqrt{2}$ characterizes an uncorrelated pair.

Minimal spanning tree

- Minimal spanning tree (MST) is used to extract the most important connections in the whole network.
- MST reduces the number of $N(N-1)/2$ pairs to only the $N-1$ most important connections while the whole system remains connected.
- How to construct MST:
  - transform the correlation matrix $\mathbb{C}$ into a distance matrix $\mathbb{D}$, discarding the diagonal elements (containing zero distances);
  - find the closest pair of assets, which creates the first two nodes in the network connected by the first link (with a weight equal to the distance d_{ij});
  - proceed to the second closest pair which creates the second pair of nodes. At this point, if a node from the second pair is already present in the network, the new node is simply connected to the existing pair;
  - proceed until $N-1$ connections are found, while the network must not be closed or create closed loops

Hierarchical tree

- MST helps us to construct hierarchical trees (HT) which are important for the analysis of clusters.
- How to construct HT:
  - determine the subdominant ultrametric distance matrix $\mathbb{D}^*$;
  - elements of $\mathbb{D}^*$ – d_{ij}^* – are defined as the maximal weight of the link which needs to be taken to move from node i to node j in the MST;
  - in matrix $\mathbb{D}^*$, we find the minimal distance d_{ij}^* and create the first pair of assets;
  - follow in connecting the assets and if more assets with same d_{ij}^* are found, the clusters are connected together

Stability of the procedure

- MST and HT might be unstable, i.e. we don't know whether the link is significant or just random (it is of course possible to construct MST and HT for completely random system)
- To deal with the problem, we use a bootstrapping technique proposed by Tumminello (2007) specifically for MST and HT analysis:
  - construct the original MST and HT;
  - construct a bootstrapped time series from the original while keeping the time series length fixed;
  - MST and HT are then constructed for the bootstrapped time series and links are recorded;
  - check whether the connection in the original MST are also present in the new MST based on bootstrapped time series;
  - the share of the bootstrapped cases where the link appears between nodes i and j will be labelled as b_{ij} with an obvious range $0 \leq b_{ij} \leq 1$

Outline

Dataset – biofuels and feedstock

Asset	Source	Type
US Ethanol	Bloomberg	Spot, FOB, anhydrous ethanol
Brazilian Ethanol	CEPEA	Anhydrous ethanol
Biodiesel (2003-08)	Bloomberg	German biodiesel, spot
Biodiesel (2008-15)	Reuters	Spot price, ARA OTC
Corn	Bloomberg	1st month futures, CBOT
Wheat	Bloomberg	1st month futures, CBOT
Sugarcane	Bloomberg	1st month futures, ICE
Sugar Beets	Bloomberg	1st month futures, LIFFE
Brazilian Sugar	CEPEA	Spot USD Price
Rapeseed Oil	Bloomberg	1st month futures
Soybean Oil	Bloomberg	1st month futures, CBOT
Sunflower Seeds	Bloomberg	1st month futures
Palm Oil	Bloomberg	1st month futures

Dataset – fossil fuels and food

Asset	Source	Type
Brent Crude Oil	Bloomberg	1st month futures, ICE
German Diesel	EIA	Retail Diesel Prices
German Gasoline	EIA	Retail Premium Gasoline
US Diesel	EIA	Retail Diesel Prices
US Gasoline	EIA	Retail Premium Gasoline
Brazilian Diesel	ANP Brazil	Weighted av. consumer price
Brazilian Gasoline	ANP Brazil	Weighted av. consumer price
Coffee	Bloomberg	Arabica, 1st month futures
Cocoa	Bloomberg	1st month futures, NYBOT
Rice	Bloomberg	1st month futures, CBOT
Oranges	Bloomberg	1st month futures

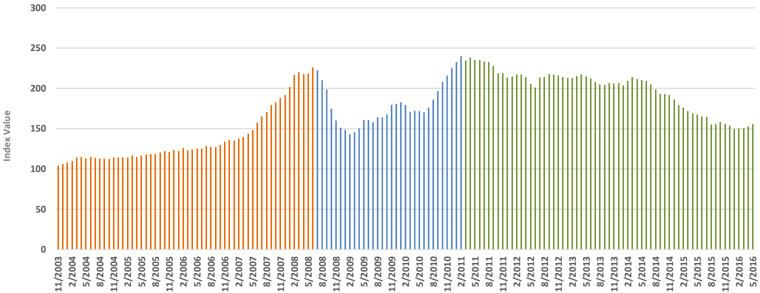
Dataset - financials

Asset	Source	Type
Dow Jones	Bloomberg	US Dow Jones Ind. Average
S&P 500	Bloomberg	US S&P 500 Index
FTSE 100	Bloomberg	British FTSE 100 Index
DAX	Bloomberg	German DAX Index
BOVESPA	Bloomberg	Brazilian BOVESPA
Federal Funds	Federal Reserve	US Fed Funds Rate
LIBOR	ECONSTATS	3 months USD LIBOR
USD/EUR	ECB	
USD/BRL	Federal Reserve	

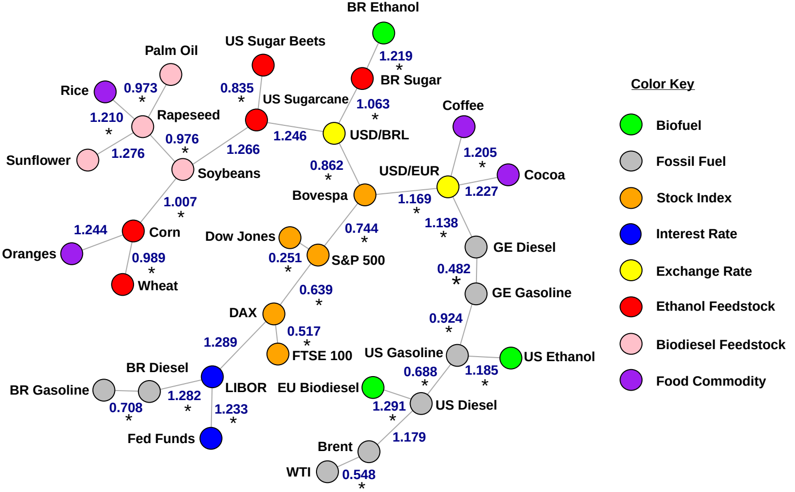
Time periods

Weekly data 2003-2016
Three subperiods, separated by peak prices on July 7, 2008 and March 7, 2011.

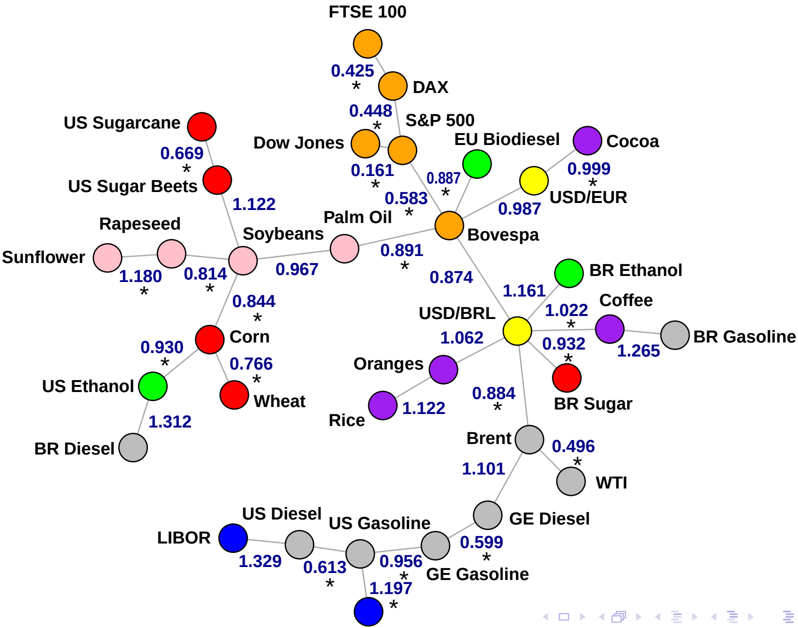
Figure: FAO Food Price Index



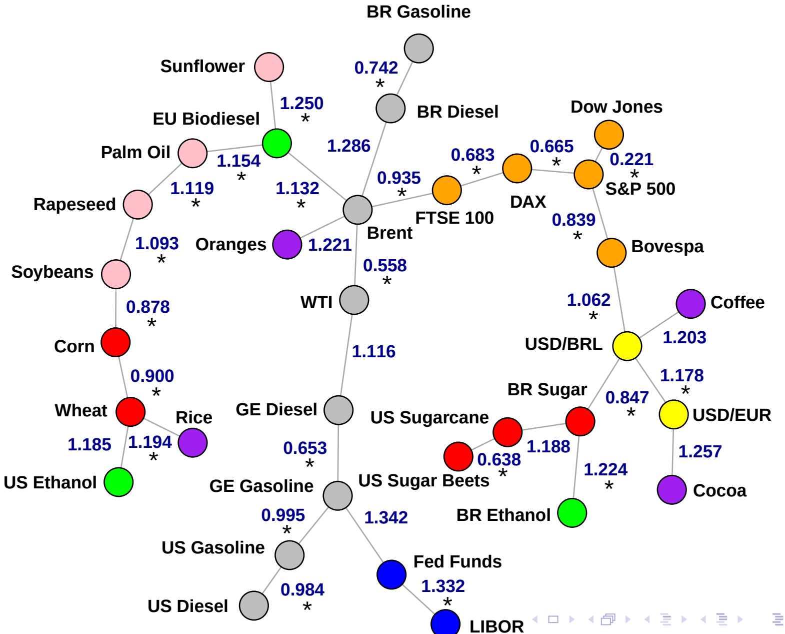
MST, 2003–2008, weekly



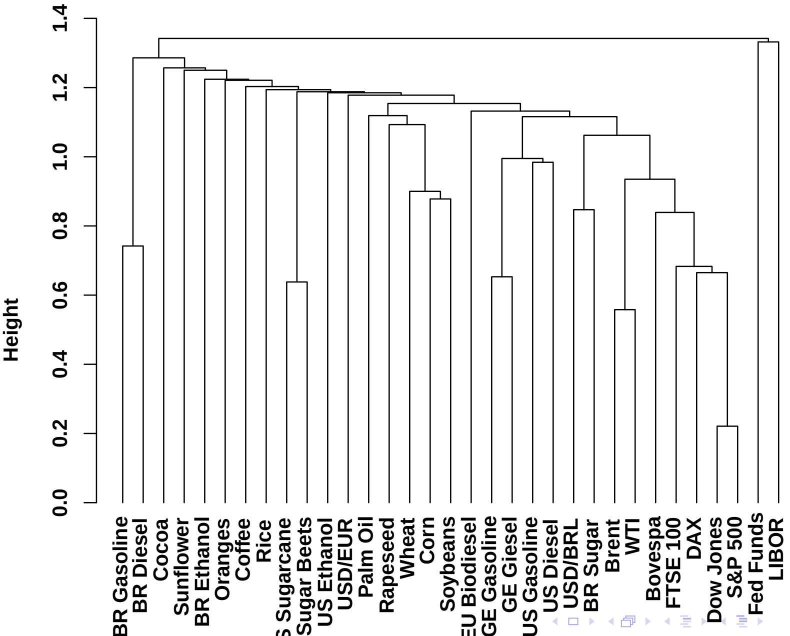
MST, 2008–2011, weekly



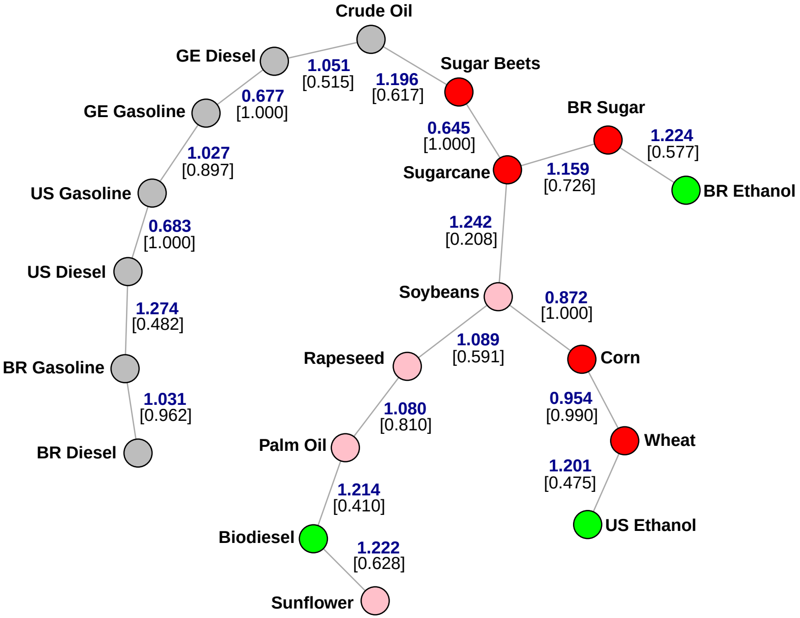
MST, 2011–2016, weekly



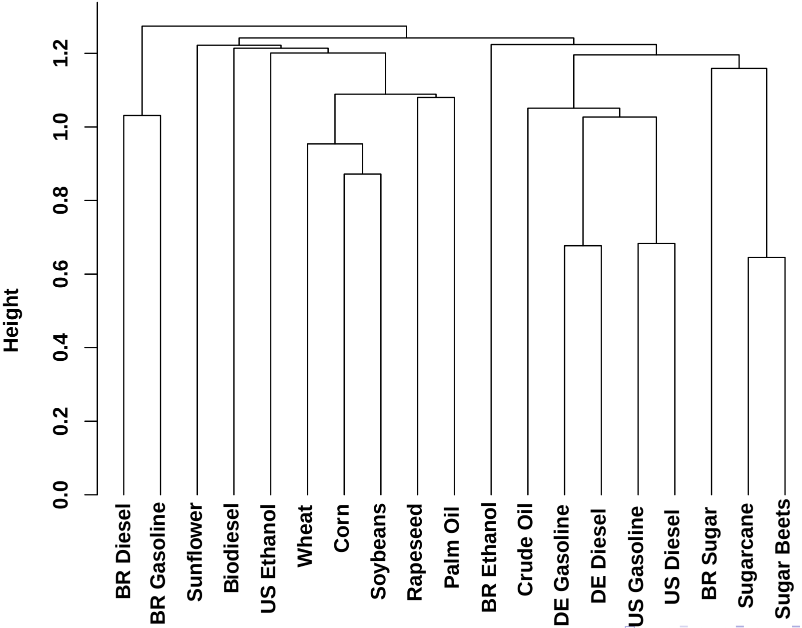
HT – All, 2011-2016, weekly



MST – No foods and financials, 2011–2015, weekly



HT – No foods and financials, 2011-2015, weekly



Summary of taxonomy results (1)

- Period 2003-2008:
  - In the short term, both biodiesel and US ethanol separated into one branch jointly with crude oil and US and EU fossil fuels. All foods (with exception of Cocoa and Coffee) separated to another branch. Brazilian ethanol connected to food branch through sugar. Brazilian fossil fuels close together, but separated from the rest of fossil fuels and crude oil.
  - In the medium term, some changes, but Brazilian ethanol still closest to sugar.
- Period 2008-2011:
  - In the short term, US ethanol and biodiesel move away from fuels cluster, towards their feedstock. UC ethanol closest to corn, biodiesel close to its feedstocks. Brazilian ethanol connected to Brazilian sugar through US/Brazil exchange rate. Brazilian fossil fuels still separated from other fossil fuels.
  - In the medium term, everything similar like in the short term. US ethanol stays closest to corn, Brazilian ethanol gets back to be closest with sugar.

Summary of taxonomy results (2)

- : Period 2011-2016:
  - In the short term, all fosil fuels (including Brazilian ones) connect into a fuels branch, with crude oils (WTI, Brent) connecting Brazilian and other fosil fuels. All biofuels keep close to a cluster of their respective feedstocks (in particular, Brazilian ethanol closest to sugar).
  - In the medium term, all biofuels continue to keep close to a cluster of their feedstocks. Brazilian ethanol as usual closest to sugar. US ethanol and biodiesel together with their feedstocks form separate branch. Brazilian fossil fuels separate away from other fuels and through oils get connected to this US ethanol/biodiesel/feedstocks branch.[These medium term results based on 2011-2015 data]

General result: at the beginning (close to 2003) biofuel prices determined by fuel prices, gradually we observe a formation of biofuel link between oil and food prices. (missing link story)

Outline

Discussion

- Structure of the network differs for the three analysed period. Are the chosen periods the right ones?
- Potential non-linear relationships between commodities
- Wavelet coherence

Outline

Wavelets – Definition

- A wavelet is a real-valued square integrable function, $\psi \in L^2(\mathbb{R})$, defined as:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-u}{s} \right) \tag{3}$$

  location parameter u determines the exact position of the wavelet
  scale parameter s defines how the wavelet is stretched or dilated.
- Wavelets have significant advantages over basic Fourier analysis when the object under study is locally non-stationary and inhomogeneous. Also we do not loose time information.

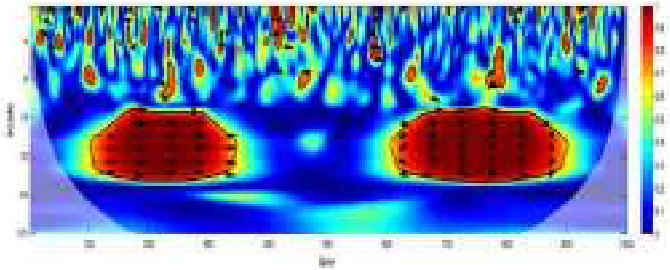
Wavelet Coherence

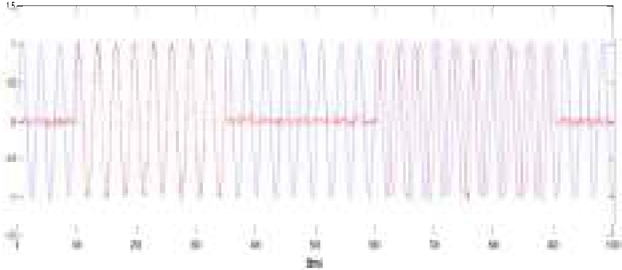
- Continuous wavelet transform $W_x(u,s)$
- $W_x(u,s) = \int_{-\infty}^{\infty} x(t) \frac{1}{\sqrt{s}} \overline{\psi\left(\frac{t-u}{s}\right)} dt,$
- represents local energy (variance) at a specific scale (frequency) u at a position s.
- Squared wavelet coherence coefficient $0 \leq R^2(u,s) \leq 1$
- $R^2(u,s) = \frac{|S(s^{-1}W_{xy}(u,s))|^2}{S(s^{-1}|W_x(u,s)|^2)S(s^{-1}|W_y(u,s)|^2)},$

Phase Difference

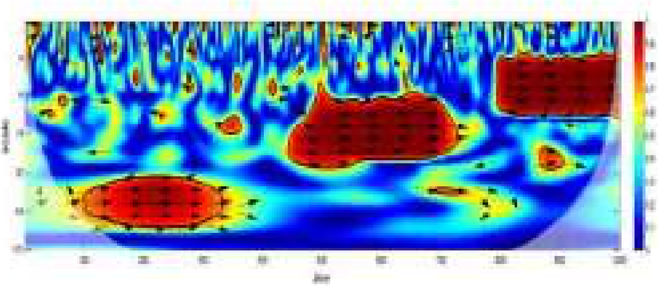
- $\varphi_{xy}(u,s) = \tan^{-1}\left(\frac{\Im\left[S(\frac{1}{s}W_{xy}(u,s))\right]}{\Re\left[S(\frac{1}{s}W_{xy}(u,s))\right]}\right)$ where $\Im$ and $\Re$ represent an imaginary and a real part operator, respectively.

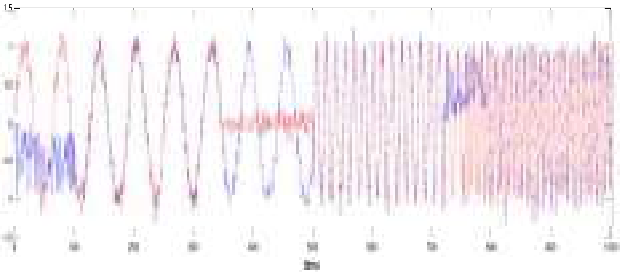
Wavelets – Intro(1)





Wavelets – Intro(2)



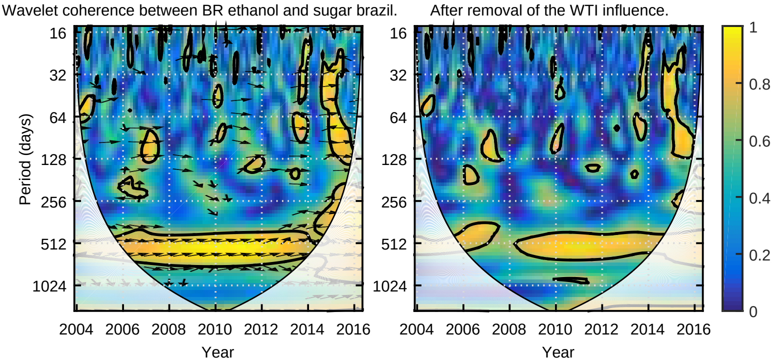


Partial Wavelet Squared Coherence

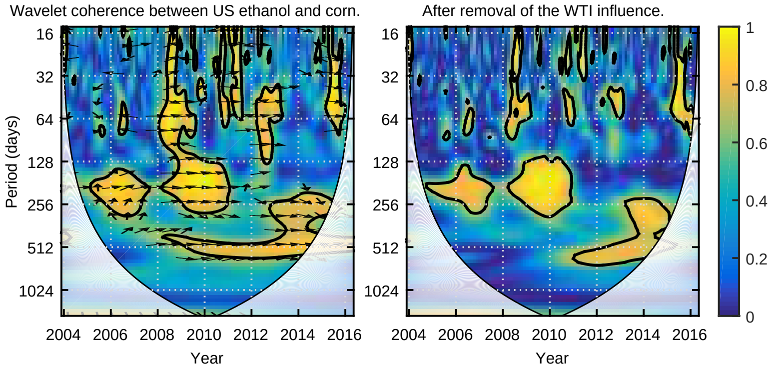
- $RP^2_{y,x_1,x_2} = \frac{|R_{yx_1} - R_{yx_2} R^*_{yx_1}|^2}{\left(1 - R^2_{yx_2}\right)^2 \left(1 - R^2_{x_2 x_1}\right)^2}$

Outline

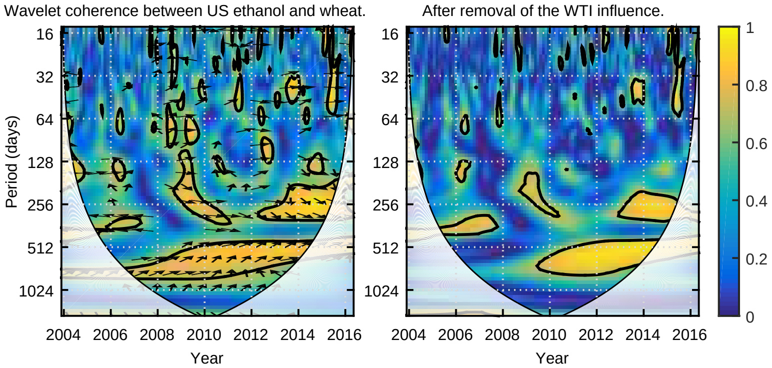
Brazilian Ethanol versus Feedstock



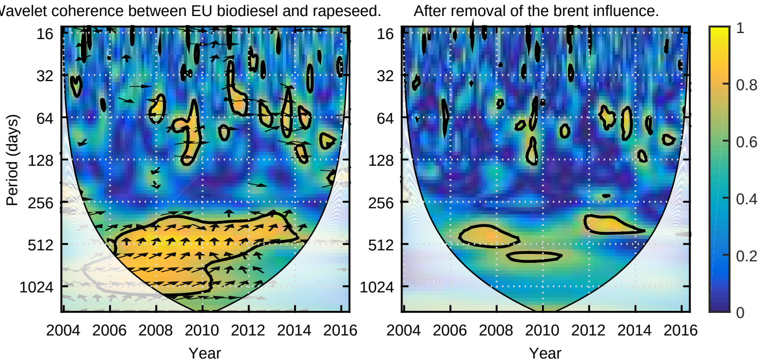
US ethanol versus feedstock



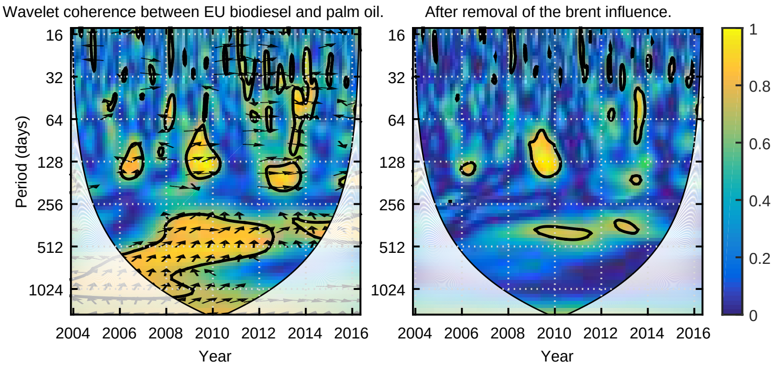
US ethanol versus feedstock



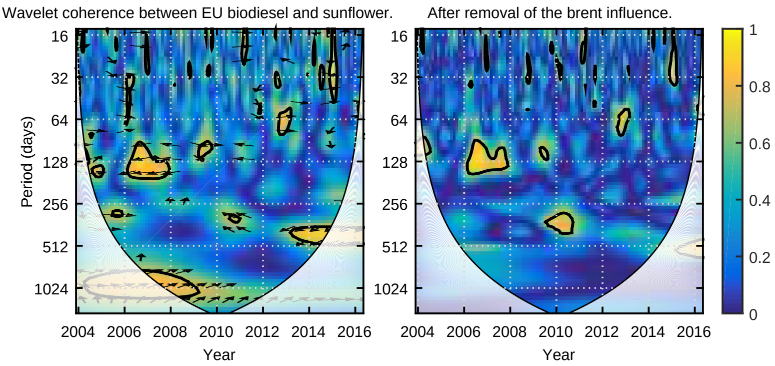
Biodiesel versus Feedstock



Biodiesel versus Feedstock



Biodiesel versus Feedstock



Biodiesel versus Brent