A Time-Frequency Analysis of Oil Price Data
The Cournot Centre and Foundation

The Cournot Centre is an independent, non-profit organization. It has created a forum where specialists meet to advance theories and increase understanding of the economic and social sciences and their epistemology. Like the Centre, the Foundation pursues catalysing work in the tradition of Augustin Cournot, accelerating theoretical formulations. Under the aegis of the Fondation de France, the Foundation puts into perspective the probabilistic paradigm shift originating in mathematics and spreading across disciplines.
A Time-Frequency Analysis
of Oil Price Data

Josselin Garnier and Knut Sølna

Prisme N° 33

October 2017

1The authors wish to thank Bob Solow and Jean-Philippe Touffut for challenging them to understand oil price dynamics; Olivier Boylaud for providing data; and Han Wang for contributing to the mathematical formulations.
Summary

Oil price data have a complicated multi-scale structure that may vary over time. We use time-frequency analysis to identify the main features of these variations and, in particular, the regime shifts. The analysis is based on a wavelet-based decomposition and a study of the associated scale spectrum. The joint estimation of the local Hurst coefficient and volatility is the key to detecting and identifying regime shifts and switches in the crude oil price since the mid-1980s until today.
## Contents

1. Introduction .................................................. 1

2. Oil Price Data ................................................. 2

3. The Scale Spectrum for Log Oil Price Data ............... 4

   4.1 Local Hurst Exponents ..................................... 9
   4.2 Local Volatilities .......................................... 10
   4.3 Spectral Misfits ........................................... 13

5. Modeling with Uncorrelated Returns ....................... 13

6. Conclusions and Perspectives .............................. 16
1 Introduction

Market price fluctuations are commonly modeled by discrete time random walks or continuous diffusion driven by standard Brownian motion. The pioneering work of Mandelbrot [16, 17] showed that such frameworks in most cases poorly describe the price fluctuations. We can refer to [18] for a historical presentation. In this text, we consider oil price data for the period from May 1987 to September 2017. The prices are recorded every trading day. We seek to understand the time-frequency character of the data. As we show below, the spectral characteristics of the data may change over time and, in particular, signify regime shifts. We describe the scale spectrum of the oil price data, which is the main tool we use to unravel the time-frequency, or time-scale, structures of the data [1, 15]. This shows that the oil price data exhibit a power-law character, in that their spectrum, as a function of the frequency, obeys a power law, or, equivalently, that its scale spectrum obeys a power law as a function of the scale. We understand temporal scales as corresponding to the periods associated with the different frequencies. This is similar to the type of scaling one sees, for instance, in turbulence data [20, 22]. What is somewhat surprising with the oil price data is that this power law persists over many scales, in fact, essentially over all the available scales. The power-law parameters we discuss here are the Hurst exponent and volatility. The Hurst parameter determines how price changes over different time intervals are correlated, and it also characterizes the scaling law of the energy in the price fluctuations over different frequencies. The volatility determines the typical magnitude of the relative price changes. The character of the power law, however, varies over time, and, as mentioned, suggests regime shifts. The variations in the power-law parameters reveal periods in the data that cannot easily be seen directly.

The estimation method for the power-law parameters of oil price data that we propose is based on wavelet decomposition. It has already been proposed in the literature to estimate the local
fractal dimension – or Hurst exponent – of different sets of data, either synthetic time series [13, 21, 28] or experimental (physical or financial) time series [4, 5, 6, 7, 24], and to use wavelet-based decomposition to do so. It has long been identified that wavelet analysis is an important addition to time-series methods with practical applications in economics and finance [25, 12]. Wavelets, for example, have been used to study the evolution of the impact of oil price changes on the macroeconomy [2], to investigate market efficiency in futures markets for oil [30], and to estimate the Hurst exponent of the crude oil price [10] (other methods have been used to estimate the Hurst index [3, 9, 26, 14]). The wavelet-based method for the joint estimation of the Hurst exponent and volatility that we propose is different, however, and the analysis of the two parameters reveals a more detailed structure. The joint estimation method that we propose is original in that we use continuous or non-decimated wavelet coefficients that are strongly correlated. The joint analysis of the two parameters allows for sharper detection and identification of regime switches.

In Section 2 we present the oil price data. In Section 3, we plot the scale spectra of the data sets and show that price modeling with a local power-law structure is indeed appropriate. We describe the structure of the Hurst exponent and volatility estimated over successive overlapping windows in Section 4. We underline that joint estimation of the two parameters is critical, in particular, for obtaining a correct assessment of the volatility variations. For comparison, we also study the estimated volatility for the standard (local geometric Brownian motion) model in Section 5, when the Hurst exponent $H$ is assumed to be equal to 1/2, corresponding to a Brownian model with independent increments or returns.

2 Oil Price Data

In Figure 1 the red dashed line shows the raw daily oil price data vector for the "West Texas Intermediate (WTI), Spot Price Free
On Board (FOB) (New York)” (hereafter “West Texas”) in dollars per barrel: $P(t)$ is based on daily price records. The solid blue line represents the “Europe Brent Spot Price FOB (London)” (hereafter “Brent”) in dollars per barrel. The daily data are available from May 1997 to September 2017 [29].

![Raw Pricing Data](image)

**Figure 1:** Full data vector for West Texas (red dashed line) and Brent (solid blue line).

In Figure 2, we show the returns process for the Brent and West Texas data sets defined by

$$R_n = \frac{P(t_n + \delta t) - P(t_n)}{P(t_n)},$$

(1)

where $P(t)$ is the “raw” pricing data and $t_n = n\delta t$, with $\delta t$ being one day.

We observe here from the magnitude of the fluctuations seen in Figure 2 that the volatility of the returns process is not constant, but rather exhibits temporal variations. We will analyse the scale contents of the data shown in Figures 1 and 2 in the next sections.
Figure 2: Returns for West Texas (red dashed line) and Brent (solid blue line).

3 The Scale Spectrum for Log Oil Price Data

We show in Figures 3–5 the scale spectra for the complete log-transformed data (Figure 3) or for various subsegments of the log-transformed data (Figures 4–5). The scales are expressed in units of years. They show the energy in the different scales, which are analysed below. We compute the scale spectrum with respect to the “Haar” wavelet basis, making it relatively robust with respect to noise. The first-level Haar coefficients correspond to the consecutive differences in the data. In our case, the data is the log prices so that the Haar coefficients form an analogue of the returns process. The Haar coefficients at higher levels correspond to differences in local averages of increasing length. Thus, the higher order differences may be thought of as returns over longer intervals.

In Figures 3–5, we can see linear behavior in the log-log plot corresponding to a power law, making it possible to discuss the power-law parameters, which are the Hurst exponent $H$ and volatility $\sigma$. The Hurst exponent characterizes the power-law decay of
the spectrum. It also characterizes the existence of correlations between the increments of the process. If $H = 1/2$, then the increments are not correlated, which is the typical case for Brownian motion or a similar process. If $H > 1/2$, the increments are positively correlated, which is the phenomenon called “persistence”. If $H < 1/2$, then the increments are negatively correlated, which is the phenomenon called “anti-persistence”. The smoothness of the process increases with $H$, since consecutive increments become more correlated as $H$ increases.

In Figure 3, we show the “global power law” for West Texas (red dashed line) and Brent (blue solid line), obtained from the complete log-transformed data. The spectra conform with a global power law, with estimated Hurst exponents $H = .46$ (Brent) and $H = .44$ (West Texas), and estimated volatilities $\sigma = .34$ (Brent) and $\sigma = .32$ (West Texas). As we will see below, this global power law is consistent with a situation in which the Hurst exponent and volatility vary over subsegments. It is striking to see that a non-trivial power law (that is, a power law with $H \neq 1/2$) emerges very clearly from the financial data, while it is difficult to exhibit such a structure from physical data (such as the distribution of energy among turbulence vortices), for which there are, on the contrary, theoretical arguments to support a power law (for instance, Kolmogorov’s theory of turbulence) [22]. In Figures 4 and 5, the spectra correspond to the first and last 16 years of the price data (the periods 1987–2002 and 2002–2017). The strength of the price fluctuations are slightly stronger in the latter half of the data for the longer scales. The estimated volatilities are $\sigma = .28$ for the first 16 years and $\sigma = .38$ for the last 16 years for the West Texas data (red dashed lines). The estimated volatilities are $\sigma = .32$ for the first 16 years and $\sigma = .38$ for the last 16 years for the Brent data (blue solid lines). The qualitative behavior of the spectrum is, however, similar for the two halves of the data. The associated Hurst exponent estimates are $H = .41$ for the first 16 years and $H = .47$ for the last 16 years for the West Texas data (red dashed lines). The solid blue lines are the corresponding spectra.
Figure 3: The “global power law” for West Texas data (red dashed line) and Brent data (blue solid line). When the scale energy is computed from the complete log-transformed data, we observe approximately a linear scale spectrum. The dotted lines represent a fitted spectrum. Here and below the spectrum is computed with Haar wavelets and is “continuous” in space and scale.
for the Brent data set with the associated Hurst exponent estimates being $H = .43$ and $H = .49$, respectively. The dotted lines in the figure are the corresponding model spectra that represent a “perfect” power law with the estimated exponents. We can observe that the spectra retain an approximate power-law behavior on very long scales. Moreover, the power law is very similar in the two halves of the data. This happens when we average the scale energies over a relatively long time period of 16 years. We will see below that if we consider the data over shorter subsegments, the overall energy of the process changes over time, which manifests itself in a vertical shift in the spectra, corresponding to changes in the local volatility. We will also see that this change happens in a coordinated fashion for the West Texas and Brent data sets.

![Price Power Law](image)

**Figure 4**: Scale spectra for the period 1987–2002.

4 Price Modeling with a Local Power-Law Spectral Structure

As indicated by the above data analysis, we find that the log of the oil price data exhibits a power-law spectral structure, but that the
parameters of the power law vary over time. We therefore model the log prices as:

\[
\log\left(\frac{P(t)}{P(0)}\right) = B_{H,\sigma}(t),
\]

where $B_{H,\sigma}$ is a random process with local power-law behavior, with $\sigma$ being the volatility and $H$ the Hurst exponent. The classic model process for $B_{H,\sigma}$ is fractional Brownian motion (with constant $H$ and $\sigma$) [19, 11]. The parameters $\sigma = \sigma_t, H = H_t$ will be modeled themselves as varying with respect to time, albeit on a scale somewhat slower than the price process itself. This type of modeling is referred to as multi-fractal or multi-fractional stochastic modeling [8, 23].

In order to identify the local power-law parameters, we decompose the data into overlapping segments of length $2^8$ points (segments of roughly one year) in the next section and estimate a homogeneous power law within each segment. The estimated power-law parameters are then attributed to the date corresponding to the center of the segment.
4.1 Local Hurst Exponents

In Figure 6, we show the estimated Hurst exponents. Recall that the Hurst exponent determines how the consecutive increments of the process are correlated with values larger than $1/2$, corresponding to positive correlation; values less than $1/2$, corresponding to negative correlation; and $1/2$, corresponding to uncorrelated returns and absence of arbitrage. In particular, if $H > 1/2$, then the market is not weakly efficient since it possesses long memory [27]. We use here again the log of the raw pricing data as shown in Figure 1, when we compute the scale spectrum with the West Texas data set (red dashed line) and the Brent data set (solid blue line). We use segments of length that represent approximately one year ($2^8$ points). Moreover, we move the center point of the segment by one day to get the time-series, Hurst-exponent estimates shown in the figure. This is now a time series with an observation for every day corresponding to the local Hurst exponent.

Figure 6: Estimated Hurst exponents $H_t$ for the West Texas data (red dashed line) and the Brent data (solid blue line).
Note that in Figure 6, the four periods with high Hurst exponent estimates are roughly 1990–1991, 1999–2000, 2008–2009, and 2014–2015. These are important phenomena exhibited by this time-frequency analysis. We comment more on this below, given that the four periods are even more visible when looking at the volatilities. Some of these periods are neither apparent in the raw price data directly, nor in the returns data (Figure 1-2), nor in the standard volatilities estimated with standard quadratic variations (which means assuming that the Hurst exponent is 1/2), as we will see below. We also remark that the Hurst exponent is partly lower in the West Texas data set than in the Brent data set, corresponding to the price fluctuations being somewhat rougher in the West Texas data than in the Brent data.

4.2 Local Volatilities

When we analyse the log-transformed oil price data, we simultaneously estimate two parameters: the Hurst exponent $H_t$ and the local volatility parameter $\sigma_t$ in Equation (2). Using the same segmentation as in Figure 6 above, we show the corresponding volatility estimate $\sigma_t$ in Figure 7. The volatility is given relative to the annual time scale. Note again that the power law should be interpreted as a local power law with a volatility that depends on time. The West Texas data set corresponds to the red dashed line and the Brent data set to the solid blue line.

As noted above, the figure clearly shows that there are four periods with relatively high volatility: roughly 1990–1991, 1999–2000, 2008–2009, and 2014–2015. These four periods can be related to four events, marked with crosses in Figures 6 and 7.
• The first cross, in August 1990, corresponds to Iraq’s invasion of Kuwait, and it initiates a period with high volatility and a high Hurst exponent.

• The second cross, in January 2000, corresponds to the peak of a period with relatively high volatility and a high Hurst exponent. This may be explained by the approach of the year 2000 and fear of the Y2K bug, which never occurred.

• The third cross, in September 2008, corresponds to the bankruptcy of Lehman Brothers, initiating a period with very high volatility and a high Hurst exponent. We can also note that the all-time high for the oil price was reached during trading on 11 July 2008.

• The fourth cross, in July 2014, corresponds to the massive liquidation of Brent- and WTI-linked derivatives by fund managers and the beginning of the price fall, initiating a period with a very high Hurst exponent and high volatility.

Note that the second period, around 2000, cannot be detected from the direct inspection of the raw price data or the returns data. Furthermore, the fourth (and last) special period appears to be unique as its Hurst exponent reaches .7, a manifestation of strong, positive correlations between the increments. The latter has never been reported in any financial data as far as we know. The results are very much the same for the Brent and West Texas data. In fact, the scale-based correlation structure between the two data sets plotted in Figure 8 reveals that the data sets are indeed strongly correlated at all scales.

Note that the qualitative properties with respect to Hurst and volatility estimates, as well as the special periods, are stable with respect to segmentation, in that they can also be identified as doubling or halving the segment lengths. Halving the segment length makes the estimates become slightly more noisy, while doubling it causes some of the features to become slightly more blurred, in particular in the case of the 2014–2015 period.
Figure 7: Local volatility estimates relative to the annual time scale $\sigma_t$ for the West Texas data (red dashed line) and the Brent data (solid blue line).

Figure 8: Scale-based correlation between the Brent and West Texas data during the period May 1987–June 2009 (blue crosses) and July 2009–September 2017 (red crosses).
4.3 Spectral Misfits

Next, we calculate a scale spectrum misfit. This is the sum of the squared difference between the empirical scale spectra and the estimated scale spectra, with the summation going over the scales. The terms in the sum are normalized by the variance of the corresponding log-spectral point. The result is plotted in Figure 9. We observe that the spectral misfit is relatively low and statistically homogeneous with respect to time. This means that the four special periods that were detected and discussed above are well-described by the multifractional model with the Hurst parameters $H_t$ and $\sigma_t$.

![Spectral Residual](image)

Figure 9: Spectral misfits for the West Texas data (red dashed line) and the Brent data (solid blue line).

5 Modeling with Uncorrelated Returns

In Figure 10, we show the estimated volatility when we condition the Hurst exponent $H$ to be $1/2$, corresponding to a Brownian model with independent increments, or returns, which is the standard model. We can observe that the 1999–2000 period does not
appear clearly in this figure, while it does in Figure 8. The 2014–2015 period appears much less dramatic, while the multi-fractal analysis reveals its unique features characterized by a very large Hurst exponent. Note also that beyond the special periods, the standard volatility experiences somewhat strong variations, while it is rather flat (around .2) in the multi-fractal analysis.

![Volatility in % per year with H=0.5](image)

Figure 10: Estimated volatilities for the West Texas data (red dashed line) and the Brent data (solid blue line) when we condition on $H = 1/2$ to enforce uncorrelated returns.

In Figure 11, we show the spectral misfit that follows when we fix $H = 1/2$. Comparing with Figure 9, we see that this enforcement means that we do a relatively poor job of capturing important structural features in the data, as the spectral misfits are relatively high; they can also vary significantly during the special periods detected and discussed above. This is even clearer where we show the variograms for the spectral misfits in Figure 12. These properties are stable with respect to segmentation. The magnitude of the spectral misfits obtained with the multi-fractal model (with a time-varying Hurst exponent) is significantly smaller, and
we can see that the spectral misfit appears to be a white noise process supporting this modeling.

![Spectral Residual H-fixed](image)

Figure 11: Spectral misfits for the West Texas data (red dashed line) and the Brent data (solid blue line) when we condition on $H = 1/2$ to enforce uncorrelated returns.
Figure 12: Variograms for the spectral misfits. The two top lines correspond to the spectral misfits obtained from the log-transformed data when enforcing uncorrelated returns. The two bottom lines correspond to the spectral misfits obtained from the log-transformed data and processed with the multi-fractal model.

6 Conclusions and Perspectives

We have analysed oil price data with a view toward identifying regime switches. We have found that a scale spectral analysis of the log prices is an efficient approach for identifying regime shifts. The time-frequency analysis involves computing the scale spectrum and fitting it to a power law, which for the log-scale spectrum corresponds to a linear model (Figures 3–5). The different scale spectral points can be associated with energy of returns at different scales. The special regimes can be associated with enhancement of both volatility and persistence (Hurst exponent) (Figures 6–7).

A striking result is that the power-law behavior for the log prices can be seen over all scales available in the data (Figures 3–
It is very interesting, however, that if one looks at subsegments of the data, one observes a “local” power law with a local Hurst exponent and local volatility that varies over time and, in particular, that announces the emergence of the special regimes (Figures 6–7). They fluctuate, however, in a coordinated fashion, and in a way so as to generate a power law also on the macroscale. We refer to the time series defined by the estimates of the Hurst exponent and volatility in the subsegments as the inferred volatility and Hurst exponent processes.

It is important to note that to generate the inferred processes, the lengths of the subsegments are fixed, and we move the window one day at a time. The choice of the length of the subwindows is guided by the effective signal-to-noise ratio: we need to choose windows (a) large enough to have enough data to be able to estimate the local power-law parameters with sufficient precision, and (b) small enough to resolve the local power-law parameters without too much bias.

From the point of view of price process modeling, it will be interesting to further our research by trying to understand how a local power law can be associated with a global power law on the macroscale and the relation between local and global parameters. Such an analysis will likely involve a multi-scale asymptotic analysis. It is also central to understanding and better quantifying the arbitrage involved with the inferred parameters of the type observed here. Will a small amount of friction, in terms of a typical transaction cost, for example, remove the possibility of arbitrage, or is intrinsic arbitrage a central ingredient of special regimes as observed here?
References


The Prisme Series

The Prisme Series is a collection of original texts that focus on contemporary theoretical issues. The authors are contributors to the Cournot series of conferences, panels and seminars.

Latest releases:

32. Why Speech Technology (almost) Works
   Mark Liberman

31. How to Flee Along a Straight Line: Tracking Self-Repelling Random Walks
   Laure Dumaz

30. The Evolving Connection between Probability and Statistics
   Noureddine El Karoui

29. How Quantum Can a Computer Be?
   Elham Kashefi

28. The Dynamics of Capitalism and Worker Participation: A Long-Term Analysis
   Bernard Gazier & Olivier Boylaud

A complete list of publications can be found at www.centre-cournot.org
Josselin Garnier is a mathematician and professor at École Polytechnique and a researcher at the Centre de Mathématiques Appliquées (CMAP). He is an expert in waves and imaging in random media and in uncertainty quantification.

Knut Sølna is a mathematician and professor at the University of California at Irvine. A recipient of the Cournot Foundation’s Senior Research Fellowship 2017/2018, he is currently a visiting professor at École Polytechnique. Sølna is an expert in waves in random media and in mathematical finance.