Utilizing Wavelet Decomposition to Trade S&P 500

Nina Gnedin, Jeongseog Lee, Vadim Munirov, Chase Perlen, and Fangying Shi

May 11th, 2016

Introduction

The question as to whether future prices are efficiently priced is of significant interest amongst both academics and practitioners. A future refers to a contract in which the buyer agrees to purchase an asset at a specified date in the future at a specified price. As traders, we hope to answer this question with a resounding no. If we can better predict future spot prices than futures, we can expect to make money by trading the discrepancy. Specifically, if our predicted price is less than the future price, we will sell the future and, conversely, if our predicted price exceeds the future price, we will buy the future. In the former case, upon expiry of the future, we can buy in the market at the lower spot rate, per our forecast, to cover our short. In the latter case, upon expiry of the future, we can sell into the market at the higher spot rate, again per our forecast.

Motivation

Motivated by Reference **3**, we initially investigate whether such methodologies can be applied to oil futures pricing. However, we feel that the Fourier transform as is, is less applicable to oil markets which are highly susceptible to strictly localized shocks, for example wars, OPEC meetings, etc. We believed the discrete wavelet transform method would be a more appropriate alternative, as it is a decomposition over the time domain which theoretically enables it to account for such localized perturbances. However, we find that such a model fails to outperform traded oil futures in predicting the future spot prices. On the other hand, when applied to the S&P 500 index, we find the discrete wavelet transform produces a successful trading strategy.

In this paper we offer a brief summary of the underlying mathematics of the discrete wavelet transform. Next, we discuss its application both, unsuccessfully, to oil futures trading and, successfully, to trading the S&P 500. We conclude with suggestions for continued research.

Wavelet decomposition

Fourier vs. Wavelet Decomposition

Fourier decomposition has been a canon of signal analysis for almost two centuries. The transform decomposes signals into combinations of complex exponentials as

$$\mathcal{F}(\xi) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i\xi x}dx\,. \tag{1}$$

The Fourier coefficients dictate what types of frequencies are mixed in different amplitudes. However, Fourier decomposition performs poorly at handing local dynamics. Fourier is also quite weak when the signal is non-stationary, i.e. when amplitudes of frequency components change in time. Specifically, it cannot measure the exact timing of specific frequencies entering signal, a key component of time series analysis. Moreover, Fourier is vulnerable to sharp noise elements. Since the Fourier transform of Dirac delta function spreads out to the entire frequency domain, localized sharp noise will contaminate the Fourier coefficients for the entire signal history.
Wavelet decomposition is a more recently developed solution to address these problems. Wavelet decomposition uses a family of a local function

$$\psi_{a,b}(x) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{x-a}{b}\right) \tag{2}$$

which consists of a set of translations and rescalings of L^2-function ψ. This ψ, usually called mother wavelet, satisfies the following constraints:

$$\int \frac{|\Psi(w)|^2}{|w|}dw < \infty \quad , \quad \int \psi(x)dx = 0, \tag{3}$$

where $\Psi(w)$ is a Fourier transform of ψ. Now we can well define the wavelet decomposition as obtaining the coefficient

$$\tilde{\psi}(a,b) = \int f(x)\psi_{a,b}(x)dx\,, \tag{4}$$

which allows us localize our decomposition by tuning the translation parameter a and scaling parameter b.

Multiscale resolution

We proceed to a discussion on scale analysis, which will be a crucial part of discrete wavelet transform for our purposes. We begin by constructing a sequence of subspaces in $L^2(\mathbb{R})$

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots\,. \tag{5}$$

There is a scale relation in this sequence of subspaces such that if any function $f(x) \subset V_k$, then $f(2x) \subset V_{k+1}$, vice versa. We also define the father wavelet, $\phi(x)$, as the scaling function which is a one parameter family of translation spanning the whole V_0,

$$\operatorname{span}(\phi, k) = \{g(x) = \sum_{k \in \mathbb{Z}, c_k \in \mathbb{R}} c_k \phi(x-k)\} = V_0\,. \tag{6}$$

From the nested sequence and scale relation before, if $\phi(x) \in V_0$, then $\phi(2x) \in V_1$. Since $V_0 \in V_1$, it should be spanned by shifts of ϕ, giving us the fundamental equation of the father wavelet,

$$\phi(x) = \sum_{k \in \mathbb{Z}} a_k \phi(2x-k)\,. \tag{7}$$

The translation family of the mother wavelet $\psi(x-k)$ should be always orthogonal to the translation family of the father wavelet $\phi(x-l)$, for any $k, l \in \mathbb{Z}$. So to speak, if $\phi(2^j x - k) \in V_j$, is the translation of the mother wavelet we always define the orthogonal complement of V_j in V_{j+1} which should be spanned by $\psi(2^j x - l)$. Let us define this orthogonal complement as W_j. Now whenever we scale the space by 2 in terms of the father wavelet, we can express any details in that magnified space by the mother wavelet. Now we are ready to dive into the discrete wavelet transform.

Discrete Wavelet Transform (DWT)

As previously discussed, $V_1 = V_0 \oplus W_0$. Hence the entire $L^2(\mathbb{R}) = V_\infty$ Hilbert space can be expressed as

$$L^2(\mathbb{R}) = V_0 \oplus W_0 \oplus W_1 \oplus W_2 \cdots \tag{8}$$

and, therefore, any function $f(x)$ can be decomposed as

$$f(x) = \sum_k \alpha_k \phi(x-k) + \sum_{j=0}^{\infty} \sum_l \beta_{j,l} \psi(2^j x - l)\,. \tag{9}$$

This equation is the discrete wavelet decomposition. We cannot magnify infinitely in practice, so we truncate at some point which can be schematically written as $f = a_n + d_1 + d_2 + \cdots + d_n$ for finite n. The most famous bases for DWT are Haar and Daubechies. The father of the Haar wavelet is the simple function $\phi(x) = \mathbb{1}(0 \leq x \leq 1)$. The mother wavelet is $\psi(x) = \phi(2x) - \phi(2x-1)$. The Haar wavelet is the simplest wavelet, but its discontinuity, making it difficult to approximate smooth functions.

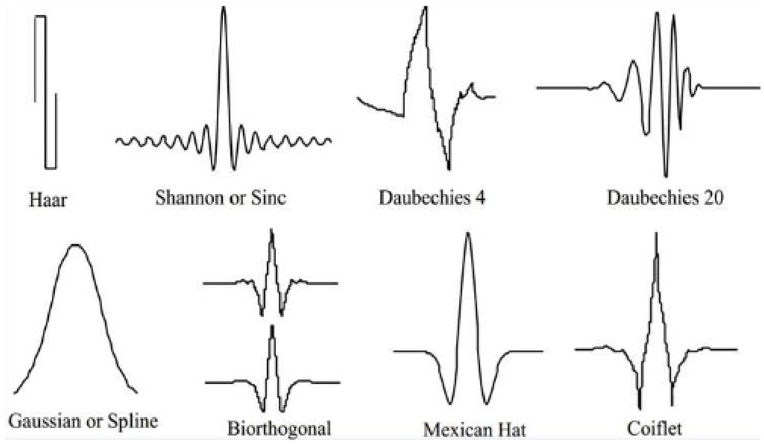
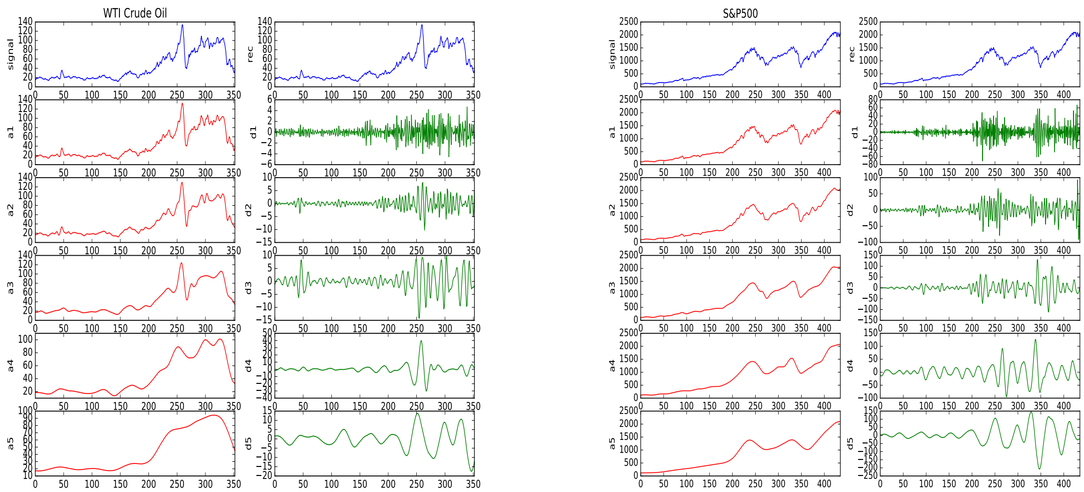


Figure 1: Examples of commonly-utilized wavelets.

Another famous basis is one invented by former Princeton mathematician Ingrid Daubechies. The Daubechies wavelet is compactly supported and characterized by a maximal number of vanishing moments. This basis is built from lowpass filters. Figure 1 provides several examples of both Daubechies and other wavelets. We will utilize the Daubechies wavelet for prediction both in oil markets and for the S&P 500.

Results

WTI Crude Oil



(a) (b)

Figure 2: Wavelet decomposition into 5 levels for (a) WTI crude oil and (b) S&P 500.

We first apply the wavelet-based methodology to the oil markets, specifically we examine monthly average WTI crude oil spot prices for the period from January 2, 1986 to March 1, 2016. However, despite the claims made in Reference [**4**] we were not successful in outperforming the futures market in predicting future spot prices.

Figure 3 shows the predicted spot oil prices versus actual realized spot oil prices alongside with the corresponding R-values for 1 and 4 months ahead, as well as the same plot with futures prices versus actual realized spot oil prices. We utilize R-value as a proxy for predictive power of our method. Figure 3 clearly suggests that current futures prices predict future realized spot oil prices better then the wavelet forecasting method, which means that we cannot successfully apply the wavelet decomposition method to oil market.

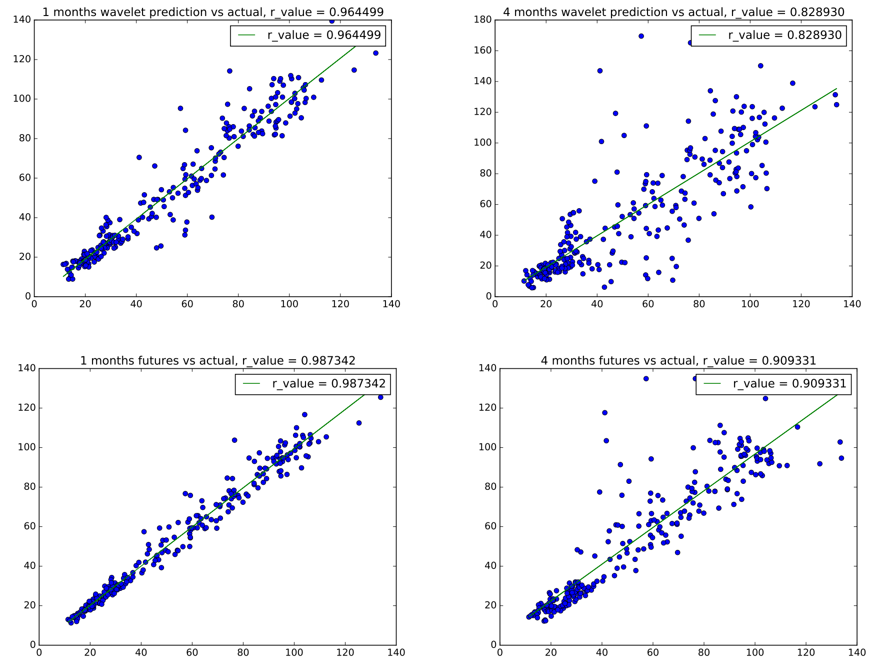


Figure 3: The wavelet predicted spot oil prices versus actual realized spot oil prices and futures prices versus actual realized spot oil prices alongside with the corresponding R-values for 1 and 4 months ahead.

S&P 500

While our methodology is not successful in oil markets, we apply wavelet forecasting to equity indices, specifically the S&P 500 where we find the results to be quite promising.

Figure 4 shows the predicted value of S&P 500 index versus actually observed S&P 500 index alongside with the corresponding R-values for 1, 2, 3 and 4 months ahead. We can see that R-values are much higher than those for the oil market suggesting that the wavelet analysis should work for S&P 500 prices.

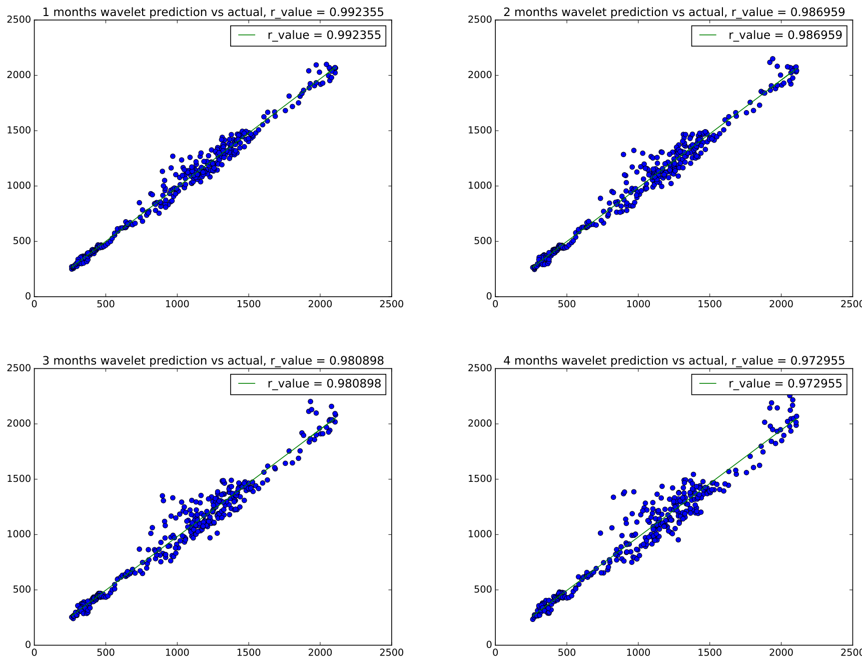


Figure 4: The wavelet predicted S&P 500 prices versus actual realized S&P 500 prices alongside with the corresponding R-values for 1, 2, 3 and 4 months ahead.

Given these promising results we construct the following trading strategy.

Trading strategy

1. *Historical prices.* We obtain S&P 500 historical monthly prices for the period from January 2, 1980 to April 1, 2016 and divide them into 333 samples of length 100 each so the start date for backtesting is April, 1988.

2. *Decomposition.* For each input sample of historical prices we perform 5 level wavelet decomposition (see Figure 2 b)
$$f = a_5 + d_5 + d_4 + d_3 + d_2 + d_1. \tag{10}$$

3. *Extrapolation.* We extrapolate 4 steps into the future. For the approximation level a_5 and the highest detail level d_5 we use spline fitting of order k, while for extrapolation of d_1, d_2, d_3, d_4 we use Fourier fitting with 10 harmonics.

4. *Reconstruction.* We obtain a wavelet-based prediction of the S&P 500 index 4 steps into the future by reconstructing the extrapolated series for a_5, d_5, d_4, d_3, d_2, d_1. Based on this prediction we forecast the expected return $r_{prediction}$.

5. *Adjustments.* We find that the expected return is sensitive to the order of spline fitting k used to extrapolate a_5 and d_5, but it is quite robust to the number of harmonics used for Fourier fitting of d_4, d_3, d_2, d_1. To mitigate this issue we repeat steps 3-4 for several values of parameter k. If different forecasts agree on the sign of the expected return, then we take their arithmetic average as the predicted return, if forecasts disagree on the sign of future returns, then we deem our forecast unreliable and do not trade at that month. To further protect us from misforecasting and potentially substantial losses we cease trading whenever the absolute value of the expected return is larger than 7%.

6. *Sizing.* The wavelet decomposition produces a trigger of when to trade and in which direction. However, it does not address the question of trade sizing. To calibrate the size of our trades to mitigate riskiness of our strategy we utilize the Kelly Criterion. The Kelly Criterion is applicable given that we allow for shorting in our base case trading strategy and place a series trades over the lifespan of the portfolio. We denote our portfolio as Π.

$$\text{Trade Size} := (\Pi_{\text{Size}} \cdot \text{ Max Drawdown}) \left(\mathbb{P}(\text{Gain}) - \frac{\mathbb{P}(\text{Loss})}{\mathcal{R}_{Gain}/\mathcal{R}_{Loss}} \right). \tag{11}$$

We make the standard assumption that future returns obey a normal distribution with the average around $r_{prediction}$ and with standard deviation equal to the historical rolling volatility. We also assume that possible gain and loss are equal. Then the probabilities of loss and gain can be easily calculated yielding

$$\text{Trade Size} = (\Pi_{\text{Size}} \cdot \text{ Max Drawdown}) \left[1 - 2N\left(-\frac{r_{prediction}}{\sigma_{historical}} \right) \right], \tag{12}$$

where N is the standard normal cumulative distribution function.

7. *Transaction costs.* To model the effect of transaction costs, we introduce a simple model with friction of 1% for each trade.

Strategy	Sharpe	Annualized return, %	Max drawdown, %	Win rate, %	No trades, %
1 month	1.10	7.86	15.16	72.20	11.11
4 months	0.43	4.06	52.56	57.31	23.72
S&P 500	0.60	7.76	51.72	62.46	-

Table 1: Summary for 1 month strategy, 4 months strategy and for S&P 500.

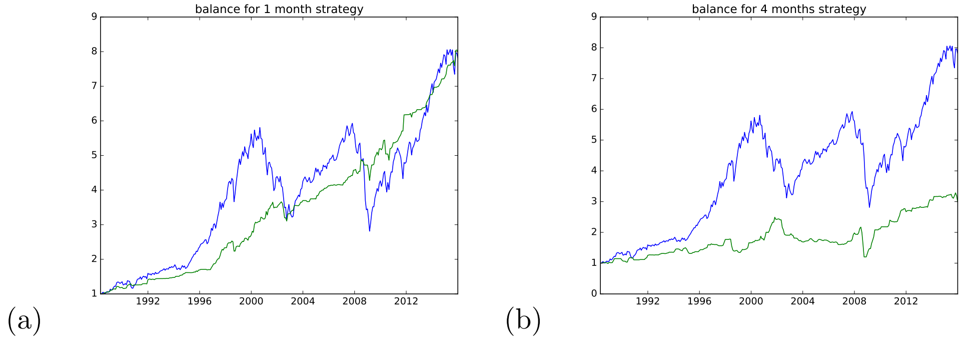


Figure 5: (a) Value of \$1 invested in 1 month wavelet forecasting strategy (green line) and in S&P 500 (blue line). (b) Value of \$1 invested in 4 months wavelet forecasting strategy (green line) and in S&P 500 (blue line).

The results of this strategy are shown in Figure 5 and Table 1, we provide summary statistics for 1 month wavelet trading strategy, 4 months wavelet trading strategy, and for passively holding the S&P 500 index. We can see that the 1 month strategy outperforms the index producing similar cumulative return, but with a Sharpe ratio almost two times higher and with the maximum drawdown of only approximately 15% instead of 56% for the benchmark. The success rate for the bets made (i.e. excluding months when we did not trade) is approximately 72%, while in 11% we do not trade because we are not confident enough in our predictions (Step 5). At the same time the 4 months trading strategy fails to outperform the index, it has only a 57% win rate and no trades (due to unreliable prediction) in about 24% of cases, which is consistent with the intuition that it is harder to predict the more distant future and, thus, predictive power diminishes as we try to forecast further into the future.

Conclusion and directions for continued research

Discrete wavelet transforms offer a solution to forecasting noisy signals susceptible to localized disruptions. We initially apply this strategy to oil markets, a classic example of both seasonality and substantial one-time shocks. However, our results suggest that the oil futures markets offer a more accurate prediction of future oil spot prices than our forecasting technique. Therefore, we test our strategy next on the S&P 500 and find that we outperform our benchmark. Our trading strategy is most successful for a shorter-term investing horizon with predictive power diminishing as our forecast window extends.

Moving forward, we can experiment with different mother wavelets. It may be the case that a different wavelet could be more suitable for a given time series. Apart from experimenting with the mother wavelets used in the transformations, we can also apply this transformation to a range of different financial time series data. To better forecast the slow-moving trend, we can also incorporate market factors into the modeling in the spirit of more traditional econometric methodologies. In this case instead of simply fitting a spline, we can build a multivariate or time-series regression model for forecasting and then add back the results of the Fourier fit forecasts.

References

1. Ramsey, J. "Wavelets in economics and finance: Past and future." *Studies in Nonlinear Dynamics and Econometrics* 6(3), 2002.

2. Thorp, E. "The Kelly Criterion in Blackjack sports betting and the stock market." *Finding the Edge: Mathematical Analysis of Casino Games*, 2000.

3. Villani, D., Ghigliazza, R., and Carmona, R. "A discrete affair." *Energy Risk*, Nov. 2003.

4. Yousefi, S., Weinreich, I., and Reinarz, D. "Wavelet-based prediction of oil prices." *Chaos, Solitons, and Fractals* 25: 265-275, 2005.