Using Wavelets-Based Time Series Forecasting to Predict Oil Prices

A Student Project

Bruce Atwood[1], Caroline Haddad[2], HelmutKnaust[3] John Merkel [4]

[1]Beloit College

[2]State University of NY at Geneseo

[3]University of Texas at El Paso

[4]Oglethorpe University

Wavelets in Undergraduate Education

Outline

Acknowledgements

- National Science Foundation
  - DUE-0442684 and DUE-0717662
- Mathematical Association of America
- Institute for Mathematics and Its Applications

Promoting Wavelets in Undergraduate Education

A hands-on approach

- Catherine Bénéteau, David Ruch, Patrick J. Van Fleet and I have a CCLI-2 grant from NSF
- Pat Van Fleet: *Discrete Wavelet Transformations* text, labs to go with the text, and created a *Wavelets Toolbox* for Matlab, Mathematica and Maple.
- Workshops
  - Workshop and materials to teach Wavelets with a hands-on approach to undergraduates.
  - Module-writing workshops to develop labs and projects for use in a course or research projects for students.

My course

- 15 weeks of three 50-minute sessions per week

My course

- 15 weeks of three 50-minute sessions per week
- Cover material through bi-orthogonal transformations

My course

- 15 weeks of three 50-minute sessions per week
- Cover material through bi-orthogonal transformations
- Instead of a final exam, there is a group project. Each group has 2-3 students.
  - Pick top 5 projects from a list or come up with an approved topic
  - They have two weeks to research topic, code solution, write a 5-page summary, present a half-hour power point presentation.

The Inspiration

The project is based on the article, *Wavelet-based Prediction of Oil Prices*, by Yousefi, Weinreich, and Reinarz (2005).

The authors propose to answer

Question: *Are futures markets efficiently priced, and can this be verified empirically?*

They answer this by using wavelets to predict oil prices.

Background

- Crude Oil trades:
  - Spot trades based on immediate delivery (*market*)
  - Future trades based on delivery at some time in the future (*NYMEX*).

Background

- Crude Oil trades:
  - Spot trades based on immediate delivery (*market*)
  - Future trades based on delivery at some time in the future (*NYMEX*).
- Spot prices and futures prices are examples of **time-series**, and the authors of the article assume it is a *stationary processe*.

Background

- Crude Oil trades:
  - Spot trades based on immediate delivery (*market*)
  - Future trades based on delivery at some time in the future (*NYMEX*).
- Spot prices and futures prices are examples of **time-series**, and the authors of the article assume it is a *stationary processe*.
  - In statistics, signal processing, econometrics and mathematical finance, a **time series** is a sequence of data points, measured at successive times.
  - **Time series forecasting** is the use of a model to predict future values based on previously observed values, e.g. oil futures.
  - A **stationary process** is a stochastic process whose joint probability distribution does not change when shifted in time or space.

Actual Problem Yousefi et al Consider

Given the spot prices from the time period Jan. 1986 - Dec. 2003, predict futures prices for 1, 2, 3, and 4 periods.

Actual Problem Yousefi et al Consider

Given the spot prices from the time period Jan. 1986 - Dec. 2003, predict futures prices for 1, 2, 3, and 4 periods.

Method Used for Predicting Prices into the Future:

- Pre-process data
- Decompose the data into 5 levels (iterations) using a Discrete Wavelet Transform.
- Extend the data 1, 2, 3, or 4 periods.
- Invert using the Inverse Wavelet transform.
- Compare the results to the actual prices and futures vs. wavelet predictions.

Pre-processing the data

- Use monthly averages instead of daily averages.

Pre-processing the data

- Use monthly averages instead of daily averages.
- De-noise the data with Donoho's Sureshrink (available in the Matlab Wavelets Toolbox). Also in VanFleet's book. (Assumes the noise is Gaussian; in general, it is not.)

Transforming the data

- In the paper, the authors claim to use the Discrete Wavelet Transform with various Daubechies filters. They claim D7 worked the best.

Transforming the data

- In the paper, the authors claim to use the Discrete Wavelet Transform with various Daubechies filters. They claim D7 worked the best.
- In reality, they must have used the *Stationary Wavelet Transform (SWT)* since the lengths of the averages and differences remained the same after each iteration.

Wavelet Transformations

- DWT_N is the *nxn* Discrete Haar Wavelet Transform (DWT or DWT1) matrix

Wavelet Transformations

- DWT_N is the *nxn* Discrete Haar Wavelet Transform (DWT or DWT1) matrix
- The filter

$$\boxed{\mathbf{h} = (h_0, h_1) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)}$$

is called the Haar filter. It is a low-pass filter.

Wavelet Transformations

- DWT_N is the *nxn* Discrete Haar Wavelet Transform (DWT or DWT1) matrix
- The filter

$$\boxed{\mathbf{h} = (h_0, h_1) = \left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)}$$

  is called the Haar filter. It is a low-pass filter.
- While

$$\boxed{\mathbf{g} = (g_0, g_1) = \left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)}$$

  is the high-pass filter.

$$DWT_N = \left[\begin{array}{c} H \\ \hline G \end{array}\right] = \left[\begin{array}{cccccccc} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 & & & 0 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & & & 0 & 0 \\ \vdots & & & & \ddots & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \hline -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 & 0 & & & 0 & 0 \\ 0 & 0 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & & & 0 & 0 \\ \vdots & & & & \ddots & & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{array}\right]$$

Example Consider a data set of $N = 8$ consecutive average monthly spot prices. These can be stored in a vector **f**.

$$\mathbf{f} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ \text{---} \\ f_5 \\ f_6 \\ f_7 \\ f_8 \end{bmatrix}$$

$$DWT_8\mathbf{f} = \frac{\sqrt{2}}{2}\left[\begin{array}{cccccccc} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{array}\right] \cdot \left[\begin{array}{c} f_1 \\ f_2 \\ f_3 \\ f_4 \\ \hline f_5 \\ f_6 \\ f_7 \\ f_8 \end{array}\right]$$

$$W_8\mathbf{f} = \frac{\sqrt{2}}{2}\begin{bmatrix} f_1 + f_2 \\ f_3 + f_4 \\ f_5 + f_6 \\ f_7 + f_8 \\ \hline f_2 - f_1 \\ f_4 - f_3 \\ f_6 - f_5 \\ f_8 - f_7 \end{bmatrix}$$

The top half, $\mathbf{a}_1$, represents the averages and the bottom half is the differences, $\mathbf{d}_1$. Both of length 4.
The process is then repeated only on the average and difference yielding $\mathbf{a}_2$ and $\mathbf{d}_2$, each half the length of $\mathbf{a}_1$.

If we use a different filter, say the Daubechies filter of length 4,or D2 (Haar is D1), then the matrix to transform our data vector, DWT2, becomes:

$$DWT2 = \left[\begin{array}{cccccccc} h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0 \\ h_1 & h_0 & 0 & 0 & 0 & 0 & h_3 & h_2 \\ \hline g_3 & g_2 & g_1 & g_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & g_3 & g_2 & g_1 & g_0 & 0 & 0 \\ 0 & 0 & 0 & 0 & g_3 & g_2 & g_1 & g_0 \\ g_1 & g_0 & 0 & 0 & 0 & 0 & g_3 & g_2 \end{array}\right]$$

One choice of low-pass filter for D2 (Daubechies filter of length 4) is

$$h_0 = \frac{1}{4\sqrt{2}}(1+\sqrt{3}) \qquad h_1 = \frac{1}{4\sqrt{2}}(3+\sqrt{3})$$
$$h_2 = \frac{1}{4\sqrt{2}}(3-\sqrt{3}) \qquad h_3 = \frac{1}{4\sqrt{2}}(1-\sqrt{3})$$

The corresponding high pass filter is

$\mathbf{g} = (g_0, g_1, g_2, g_3) = (h_3, -h_2, h_1, -h_0)$, so that DWT2 is an orthogonal wavelet matrix.

The authors experimented with Daubechies filters, and claimed that D7 (length 14) yielded the best predictions.

If the *Stationary Wavelet Transform* is used, then the data is not down-sampled, meaning the matrix required to process 8-vector **f** , is 16x8.

$$SWT_{16x8} = \begin{bmatrix} H_S \\ \hline G_S \end{bmatrix}$$

where

$$H_S = \frac{\sqrt{2}}{2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

$$H_S\mathbf{f} = \frac{\sqrt{2}}{2}\begin{bmatrix} f_8 + f_1 \\ f_1 + f_2 \\ f_2 + f_3 \\ f_3 + f_4 \\ f_4 + f_5 \\ f_5 + f_6 \\ f_6 + f_7 \\ f_7 + f_8 \end{bmatrix} \text{ and } G_S\mathbf{f} = \frac{\sqrt{2}}{2}\begin{bmatrix} -f_8 + f_1 \\ -f_1 + f_2 \\ -f_2 + f_3 \\ -f_3 + f_4 \\ -f_4 + f_5 \\ -f_5 + f_6 \\ -f_6 + f_7 \\ -f_7 + f_8 \end{bmatrix}$$

Thus each new average and difference portion has the same length as the original data set, and after 5 iterations, we have

$$\mathbf{Tr_5}(\mathbf{f}) = \mathbf{a}_5 + \mathbf{d}_5 + \mathbf{d}_4 + \ldots \mathbf{d}_1$$

Extension of the transformed data

- Each of $\mathbf{a}_5, \mathbf{d}_5, \mathbf{d}_4, \ldots, \mathbf{d}_1$ is of length N = 5.

Extension of the transformed data

- Each of $\mathbf{a}_5, \mathbf{d}_5, \mathbf{d}_4, \ldots, \mathbf{d}_1$ is of length N = 5.
- The average portion, $\mathbf{a}_5$, was extended either 1, 2, 3, or 4 periods(months) using a spline fit to the data.

Extension of the transformed data

- Each of $\mathbf{a}_5, \mathbf{d}_5, \mathbf{d}_4, \ldots, \mathbf{d}_1$ is of length N = 5.
- The average portion, $\mathbf{a}_5$, was extended either 1, 2, 3, or 4 periods(months) using a spline fit to the data.
- The authors believed that the differences or fluctuations in the data were due to seasonal changes, and therefore periodic. These data $(\mathbf{d}_1, \ldots, \mathbf{d}_5)$ were fit with a sine curve, and then extended the same number of periods as the averages.

Inverting the Transform

- We assume that the extended data was inverted using ISWT in the Wavelets toolbox. This gives us the extended data.

Inverting the Transform

- We assume that the extended data was inverted using ISWT in the Wavelets toolbox. This gives us the extended data.
- This data is now extended 1, 2, 3, or 4 periods into the future. This may now be compared to the actual price for that time period.

Analysis of the Data

- A regression analysis was performed between the actual vs. predicted data, and between the actual and the futures prices.

Analysis of the Data

- A regression analysis was performed between the actual vs. predicted data, and between the actual and the futures prices.

- The authors found that using data samples of length 100 or greater worked best. Using D7 (14-length Daubechies filters) worked best, and that the correlation of their data was typically was much higher than that of the futures. In fact, projecting 4 months ahead the correlation of the predicted data to the actual price was .98887 while the correlation to the futures data was only .773. (Frankly we are somewhat skeptical of this since we could not replicate these results...)

Problems We Encountered with the Paper

- The authors did not mention their assumption that the data is stationary, nor that they used the *Stationary Wavelet Transform (SWT)* . We had to assume this based on the graphs of their results. *Discrete Wavelet Transform(DWT)*. Also the figures were mislabeled.

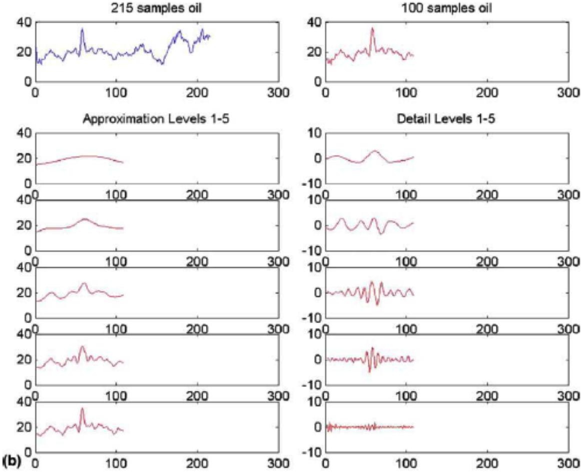
Problems We Encountered with the Paper

- The authors did not mention their assumption that the data is stationary, nor that they used the *Stationary Wavelet Transform (SWT)* . We had to assume this based on the graphs of their results. *Discrete Wavelet Transform(DWT)*. Also the figures were mislabeled.
- The SWT requires the length of the signal/data to be divisible by 2^k. Their data is not. Matlab can fill in missing data, but if that happened, then the results should be skewed. Also, it should not be possible to perform the SWT using D7 without serious wrapping issues at the ends (the end is what we are trying to extrapolate!).

Problems We Encountered with the Paper

- The authors did not mention their assumption that the data is stationary, nor that they used the *Stationary Wavelet Transform (SWT)* . We had to assume this based on the graphs of their results. *Discrete Wavelet Transform(DWT)*. Also the figures were mislabeled.
- The SWT requires the length of the signal/data to be divisible by 2^k. Their data is not. Matlab can fill in missing data, but if that happened, then the results should be skewed. Also, it should not be possible to perform the SWT using D7 without serious wrapping issues at the ends (the end is what we are trying to extrapolate!).
- Neither my students, nor our group at the module writing workshop, could reproduce what Yousefi et al did, even with access to the *Matlab Wavelets Toolbox*. The averages grew for us, while they did not for the authors.

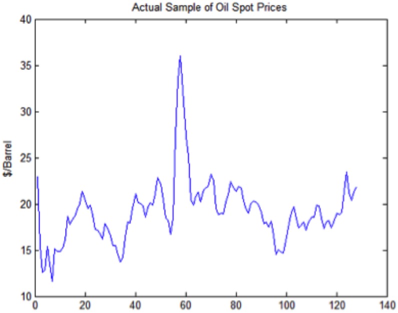
Transformed Data from Paper

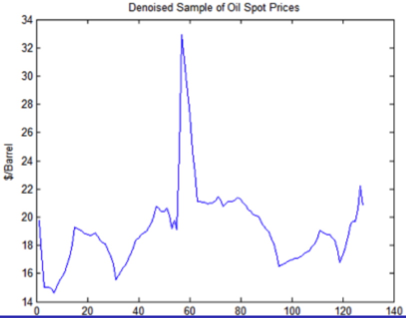


What My Students Did

Pre- Processing of Data

The data they used spanned 1986 - 2007. They chose 133 random (consecutive) samples of fixed length 128.
vspace*.25in They wrote code to de-noise the data using *Sureshrink* algorithm. It yielded better results by smoothing out some of the more extreme market fluctuations.





What My Students Did

They used the DHWT to decompose the data for 6 iterations.

What My Students Did

Extension and Processing of Data

The data was then extended.

A spline fit was used on the $\mathbf{a}_5$ and a sine fit was used on $\mathbf{d}_1, \ldots, \mathbf{d}_5$. $\mathbf{d}_6$ had only two points, so they used a weighted average to extend it.

Each was extended 4 periods (months), and then analyzed.

What My Students Did

Data Tested

They applied this to predict GE Stock Prices, Oil Prices, S and P 500 Prices, and 10-year notes.

Note: Their best results came with the S and P 500 using the DHWT.

What My Students Did

Analysis of Data

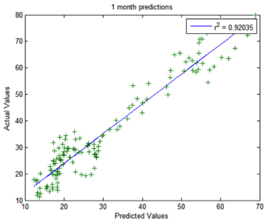
They then performed regression analysis by comparing their results to the market results. Since no one could replicate the data from the paper, they could not compare to those derived from the paper.

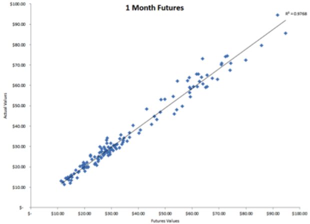
Table 1: R-Squared (Correlation Coefficient), Predictions (Futures) vs. Actual Spot Prices

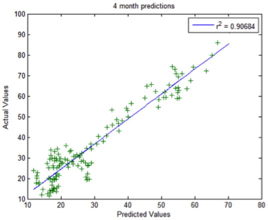
	One Month	Two Month	Three Month	Four Month
Oil Futures	0.9768	0.9506	0.927	0.9063
Oil	0.9204	0.9196	0.9088	0.9068
S&P 500	0.9707	0.9679	0.9646	0.9641
General Electric	0.8128	0.8062	0.7915	0.7834
10 Year Notes	0.6815	0.6660	0.6515	0.6441

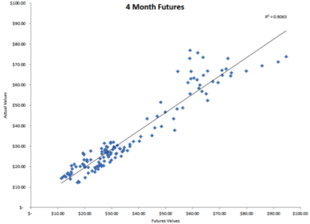
What My Students Did

Graphs of regression lines.









Summary

Possible Project Idea

- Pre-process data by appropriate de-noising
- Decompose the data into 5 levels (iterations) using the *Stationary Discrete Wavelet Transform (SDWT)*.
- Extend the data 1, 2, 3, or 4 periods.
- Invert using the *Inverse SDWT*.
- Compare the results to the actual spot prices, and futures predicted by NYMEX.

References

- Matlab Wavelets Toolbox Documentation.
  `http://www.mathworks.com/products/wavelet/`
- Van Fleet, P.J. *Discrete Wavelet Transformations*, Wiley-Interscience, 2008.
- Yousefi, S., Weinreich, I., and Reinharz, D. Wavelet-based Prediction of Oil Prices, *J. Chaos, Solitions, and Fractals* (pp. 265–275). Elsevier, 2005.

More information

- Haddad, C. Contact info: *haddad@geneseo.edu*.
- Crude Oil Spot Prices `http://www.eia.gov`
- Van Fleet, P. *Wavelets at University of St. Thomas*. `http://cam.mathlab.stthomas.edu/wavelets/index.php`.
- *Wavelets in Undergraduate Education Website* `http://math.mscd.edu/metadot/index.pl?iid=2553`.
- Van Fleet, P.J. *Discrete Wavelet Toolbox*, `http://cam.mathlab.stthomas.edu/wavelets/packages.php`.