Exploring Wavelet Methods for Time Series Forecasting

Sarah Charlton & Vince Tandaw

Johns Hopkins University

scharlt2@jhu.edu, Jtandaw1@jhu.edu

# Abstract

Wavelet decomposition is a common mathematical technique used in many signals processing applications. We apply this method to a time series of monthly oil prices, transforming the data before modelling and forecasting the series using ARIMA methods. We then compare the wavelet-based forecasts to traditional, standard ARIMA and VAR methods to determine if wavelet decomposition could provide better accuracy in forecasting. We find that our wavelet decomposition model performs slightly better than other models, yielding a lower root mean square error than the baseline models. Our work is inspired by Schlüter and Deuschle (2010), who found wavelet forecasting methods performed better than baseline ARIMA models for forecasting daily prices and daily exchange rates. We further extend that previous work by examining wavelet-based forecasting on a different time frequency (monthly vs daily), and by including VAR models as another traditional baseline for model and forecast comparison. Our findings agree with the previous work, and we conclude that wavelet-based forecasting is a useful tool.

# Introduction

Forecasting future values of an observed time series is a task with many diverse application areas, such as economics, finance, epidemiology, environmental science, and many others. Researchers and practitioners are always searching for more effective forecasting techniques. Inspired by the work of Schlüter and Deuschle (2010), in this report, we investigate wavelet methods for time series forecasting. Our research question is two-part. Firstly, do wavelet methods, in fact, produce better forecasts? Secondly, is the wavelet forecast improvement to a degree that justifies the added complexity to the time series analysis?

## Forecasting with Wavelets

We assume the reader is already familiar with standard time series analysis and time series forecasting techniques, such as autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), and exponential smoothing models; thus, we will not review those technical details here. For further reading on forecasting methods, see Diebold, F.X.’s *Elements of forecasting*.

A complete review of wavelet theory and methods is beyond the scope of this paper. Instead, we provide a brief conceptual introduction; for a more rigorous overview, see Mallat (1989) for examples. What is a wavelet? Basically, it is a little function that oscillates. Wavelet transforms are a method to represent a larger function as series of these small oscillations, it is a tool used in many areas of science and engineering, particularly in signals analysis. Below are diagrams illustrating the basic concept of wavelet multiresolution analysis (MRA).

In Figure 1, the original signal (it could be a 1-dimentional signal or a higher dimensional signal) is denoted as **S***.* For the first level of decomposition, two new series are produced, each with length equal to ½ of the length of the original series, these are denoted as **D1** and **A1**. To perform additional levels of decomposition, only the **A** series is considered. The two series produced at this level, **D2** and **A2**, are again, ½ the length of their parent, or now ¼ the length of the original series **S**. Decompositions can continue down as many levels as there is data available to split. A concrete example of this decomposition using a time series of oil prices is shown in Figure 2. For completeness, we note that the resulting components of the MRA, (**A1**, **D1**, **A2**, **D2**, …,), can be reconstituted back into the original series **S** by an inverse wavelet transform; for details see Mallat’s 1989 paper on the subject.

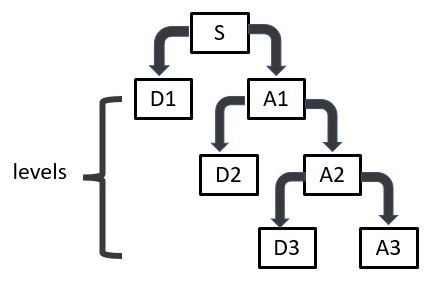


Figure 1: Illustration of a 3-level wavelet MRA decomposition of the originalsignal **S***.* Chart reproduced from S.G. Mallat’s 1989 paper.

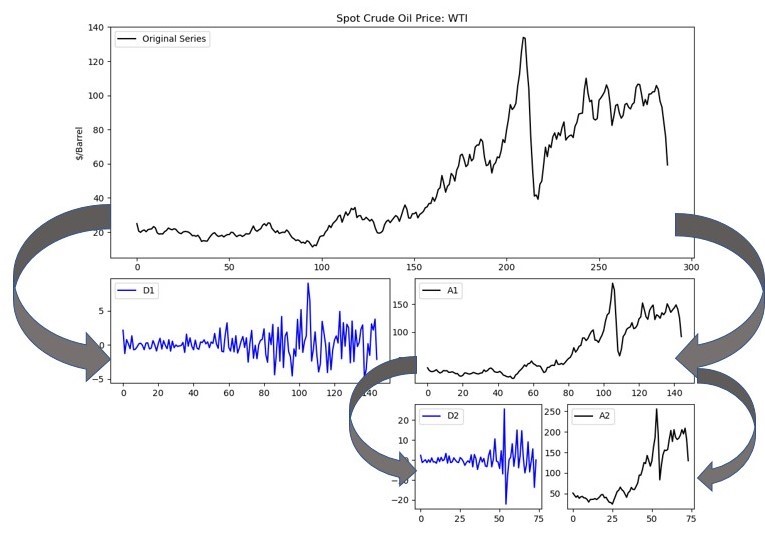


Figure 2: Two-level wavelet MRA decomposition of monthly oil prices.

Now with a basic understanding of wavelet MRA, we discuss how this technique might be useful in forecasting. Schlüter and Deuschle (2010) compare three forecasting approaches:

1. The baseline: a traditional forecasting model such as ARIMA, on the original level series.
2. Forecast the smoothed series. One general application of wavelet decomposition is for smoothing or denoising signals, so, first we can use wavelets to smooth the level series and then apply traditional forecast models to the smoothed level series.
3. Forecast the wavelet decompositions. To do method #2 above, we first break down the original level series into its multiresolution decomposition. We can forecast those individual components, and then use the forecasts to reconstruct a forecast of the original level series. This is by far the most complicated approach because it will involve forecasting multiple series.

Notional illustrations of these three approaches are pictured in Figures 3—5.

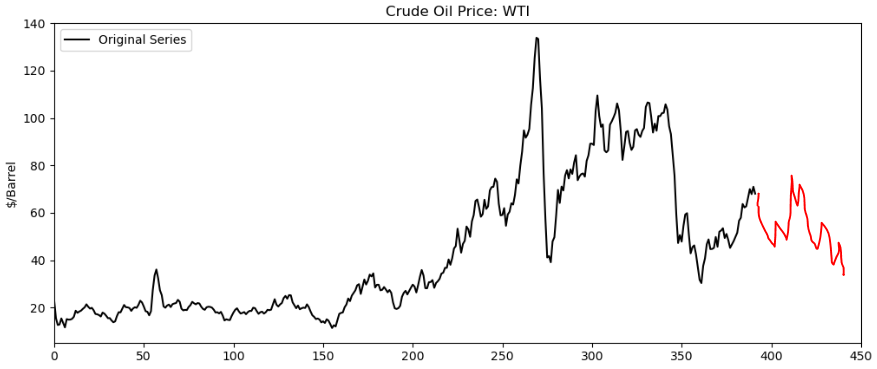


Figure 3: Traditional forecasting model on original level series (forecasting the black series).

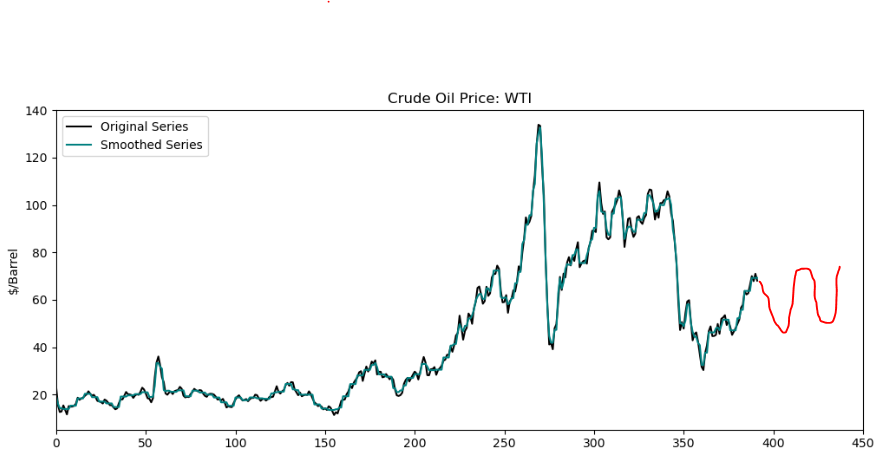


Figure 4: Traditional forecasting model on the wavelet-smoothed series (forecasting the green series).

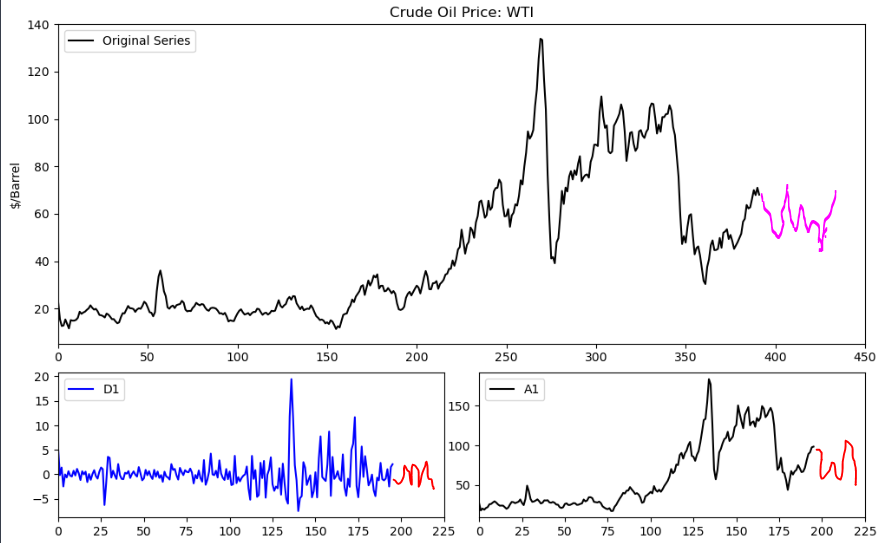


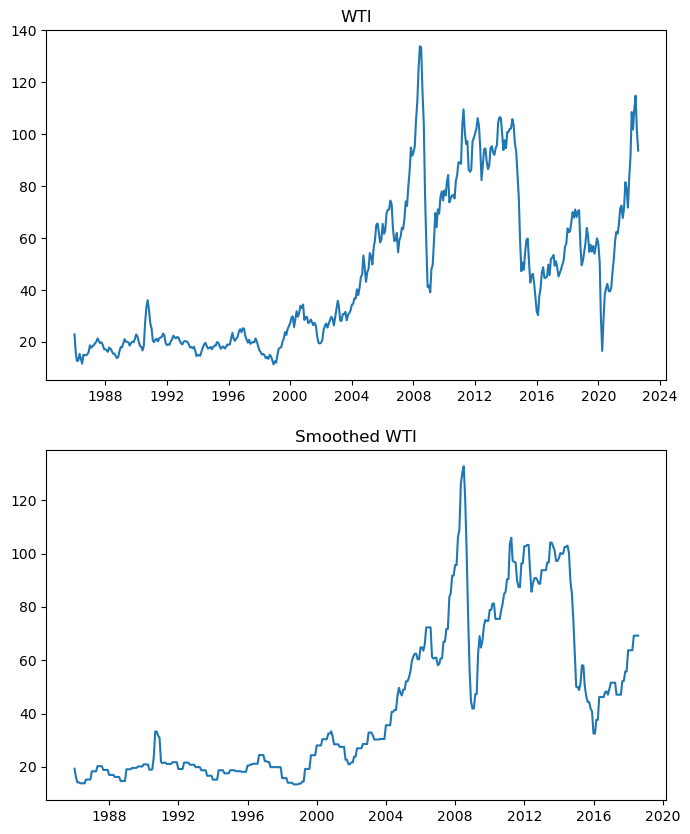
Figure 5: Traditional forecasting models on the wavelet decomposition components. The series D1 and A1 are each forecasted for h/2 forecast horizon steps (illustrated by the red series in the 2 lower graphs). These component forecasts are transformed via an inverse wavelet transform into h horizon step forecasts at the level series (illustrated by the pink series in the top graph.)

In Schlüter and Deuschle’s 2010 paper, the authors ran several experiments to determine the effectiveness of wavelet methods for forecasting. They tested 1-day and 1-week (i.e. 5-trading day) forecast horizons on four different daily time series. Using root mean squared error (RMSE) and mean absolute error (MAE), results were a bit mixed, but generally, for the 1-day forecast, ARIMA models on the MRA (approach #3) performed best among the three approaches, and for the 1-week forecast, ARIMA models on the smoothed series (approach #2) performed the best. However, the authors did not perform any diagnostics of the model fits or forecast residuals. Nor did they test whether the resulting forecasts were statistically distinguishable from one another, so it could be the case that the “better” forecast is not actually different than the traditional ARIMA model forecast. In such a case, for parsimony, we would favor the straightforward method to forecast the level series directly.

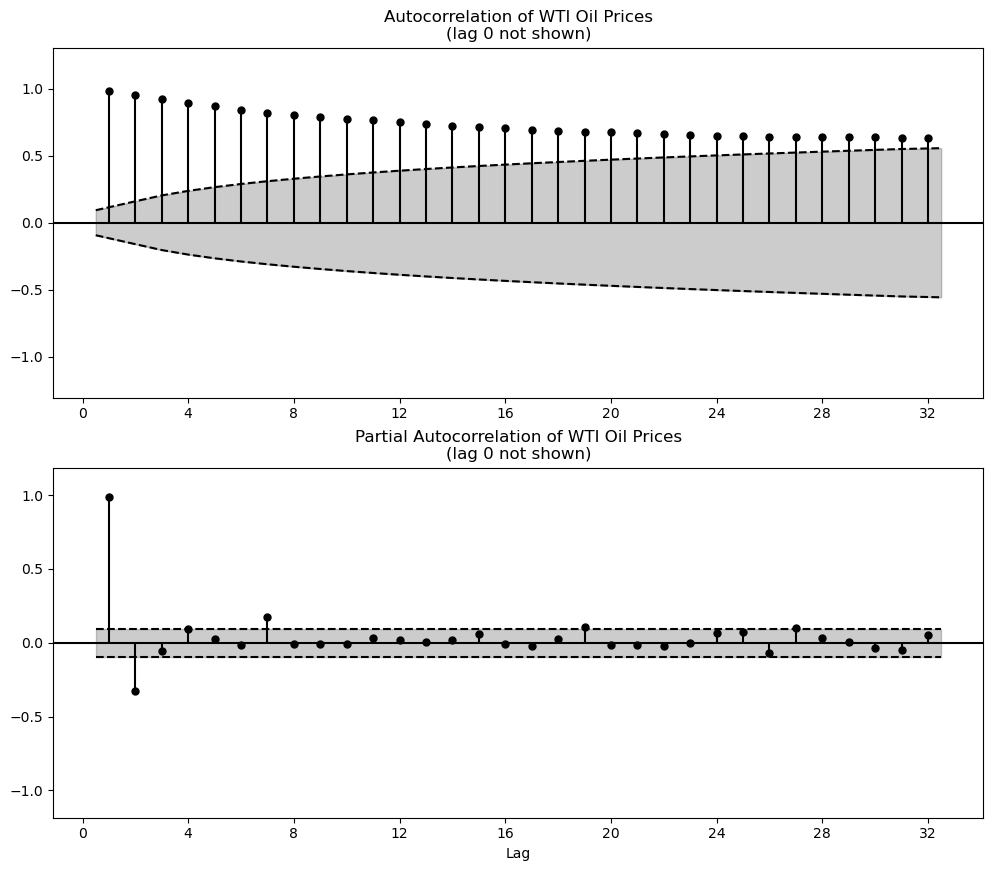
# Data

Schlüter and Deuschle used daily oil prices as one of their example series. In this work, we try forecasting monthly oil prices. We chose monthly prices for two reasons: one, to be able to demonstrate the techniques for forecasting low-frequency macroeconomic series, such as inclusion of seasonal dummy variables, and two, to explore whether the authors' conclusions about the usefulness of wavelet techniques also hold for monthly series.

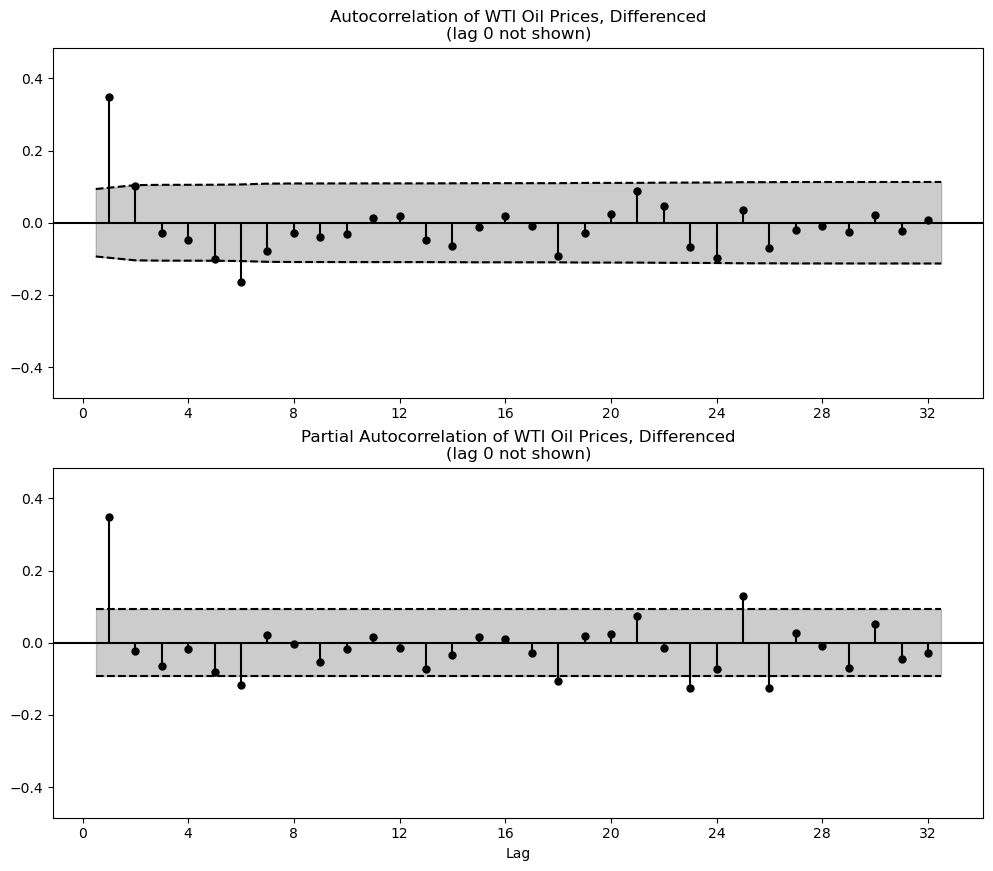
We obtain monthly crude oil prices, West Texas Intermediate (WTI) series MCOILWTICO, for the period January 1986 to March 2023. The daily oil prices used in Schlüter and Deuschle’s 2010 paper are simply identified as “WTI,” without a reference or a link to the dataset, but given that “WTI” oil prices are a common example for economic time series analysis, we believe that we are using the same underlaying data. Below graphs plot the oil prices in levels and the correlograms of the series in levels.



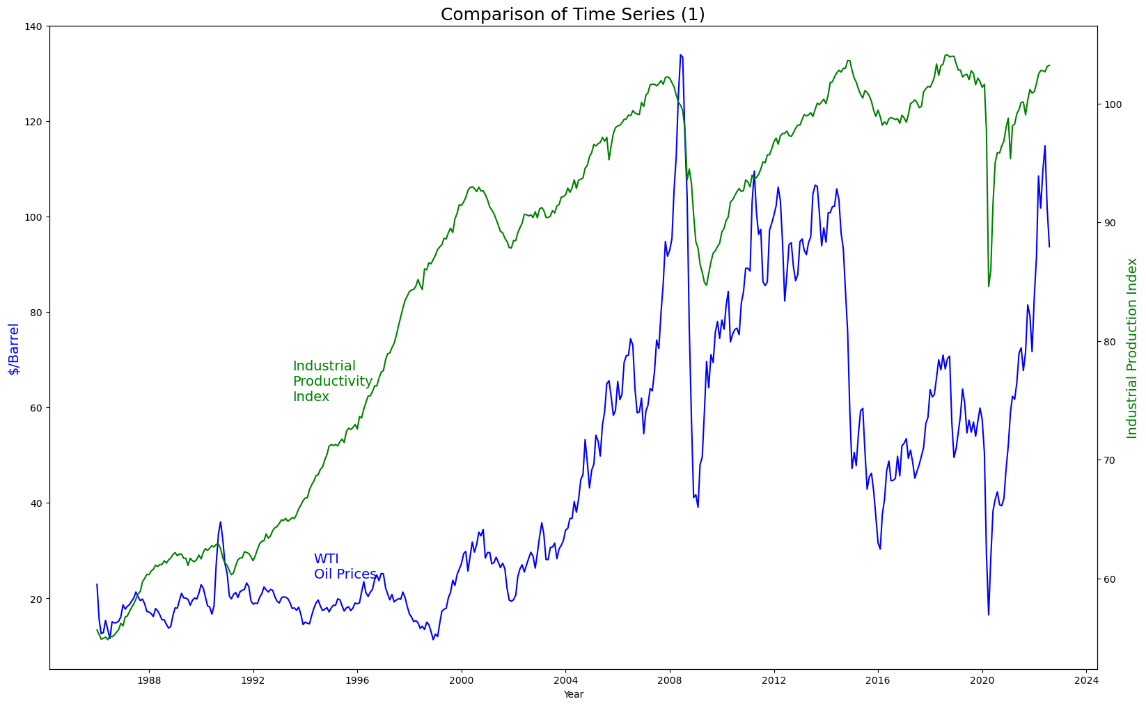
We see large spikes in 2008 and another one starting in 2020, corresponding to recent economic crises in the Global Financial Crisis and the coronavirus pandemic. This series is clearly nonstationary, as its values are quite correlated with time. This is made clear by the correlograms of the series. We can see that the slowly decaying autocorrelation does not eventually fall beneath the threshold, and the quickly decaying, sinusoidal and oscillating partial autocorrelation hint at both autoregressive and moving average effects. The correlograms are produced below:

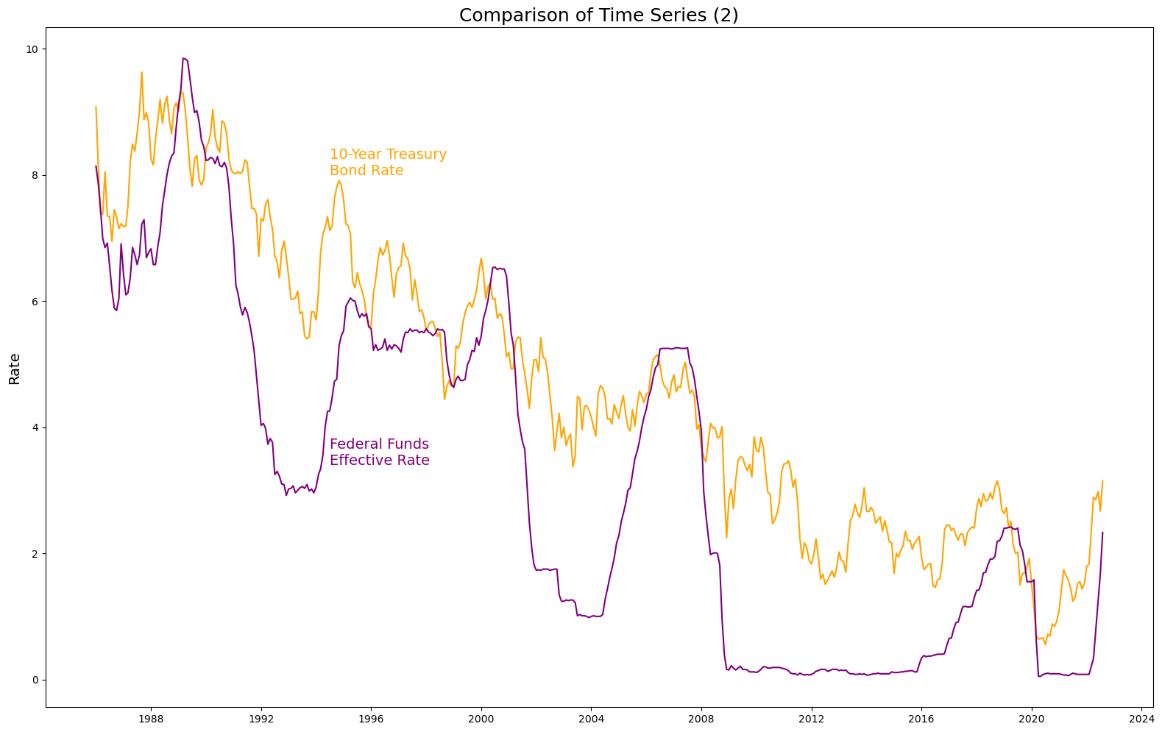


This failure to damp in the ACF plot of the series is a strong indicator that the series has a unit root, or that it is an integrated series. We can check the correlograms of the series’ first difference, produced below, and see that the series quickly decays, alleviating this problem. Furthermore, the Augmented Dickey-Fuller test strongly rejects the null of the presence of a unit root, with the test statistic of -9 and p-value being practically zero. All this evidence points to the conclusion that for forecasting oil prices, an ARIMA(p,d,q) model with *d=1* is an appropriate model choice.



As we will cover later, we also include three other time series that may help in our modeling: the Industrial Production Index, the 10-Year Treasury Bill Rate, and the Federal Funds Effective Rate. The first variable, known officially as the Industrial Production and Capacity Utilization, covers manufacturing, mining, and electric and gas utilities, which, together with construction, constitute a great portion of variation of United States national output (FRED). The second variable underpins most financial analyses as the baseline “risk-free rate”, which makes it crucial to monitor (FRED). And the third variable measures the interest rate set by the Federal Reserve at which banks borrow and sell reserves at the Fed from one another (FRED). One attractive series to also use in this analysis is the CPI, which tracks inflation through the prices of a representative “basket of goods”. However, we decided not to use CPI in our forecast model due to potential high multicollinearity problems that may arise, given that oil prices affect virtually most of the other prices in the basket, from food to fuel prices. Therefore, the inclusion of such a series may not provide much value. In a similar methodology as above, we find that these three series contained unit roots and thus were differenced once, which we found to remove the non-stationarity sufficiently, before further analysis. We plot the three series in levels below for visualization purposes:

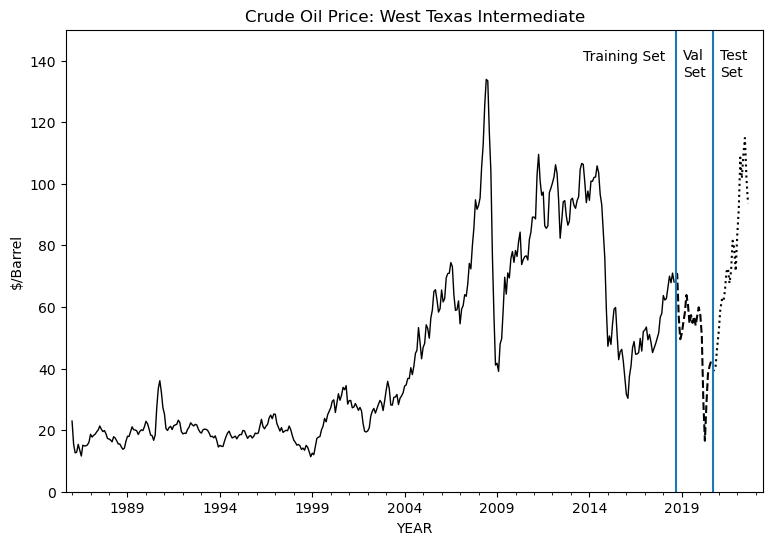




# Methodology

For wavelet analysis, it is simpler to work with series that have lengths that are a multiple of a power of 2; when series are not divisible by a power of 2, it is necessary to pad out the endpoints of the series (e.g. by reflecting the last *k* points of the series, or simply appending *k* repetitions of constant value *c*). To avoid inadvertently forecasting based on padded values of a series, we break our oil price series into sections divisible by 8.

We will use the first 392 observations to train models. The next section is the validation set, which we use to evaluate candidate models of each of the three approaches. Using the validation set, we select the "winning" forecasting model in each approach using root mean square error as our deciding metric, although the AIC and BIC will also be considered. Finally, we use the last section as the testing set to determine the ultimate forecasting "winner" from amongst the three validation set winners. The validation and testing sets have been chosen to be length of 24. This corresponds to a 2-year forecast horizon for the monthly data, which seems a reasonable forecast horizon to evaluate candidate forecasting models. Also, this forecasting horizon conveniently divides by 8, which keeps our wavelet analysis simpler. In summary, the training set corresponds to dates January 1986 through August 2018, the validation set covers September 2018 through August 2020, and the validation set covers months September 2020 through August 2022, as can be seen in the figure below.



We evaluate models based on the root mean squared error (RMSE), calculated as:

RMSE =

where are the forecasted values and are the true observed values of the time series over the forecast of *n* steps. Smaller RMSE is better. We perform all the analysis in the python programming language (Python Software Foundation), specifically utilizing PyWavelets (Lee et al.) for wavelet decomposition and Statsmodels (Statsmodels) for time series forecasting. All code is available from the authors by request.

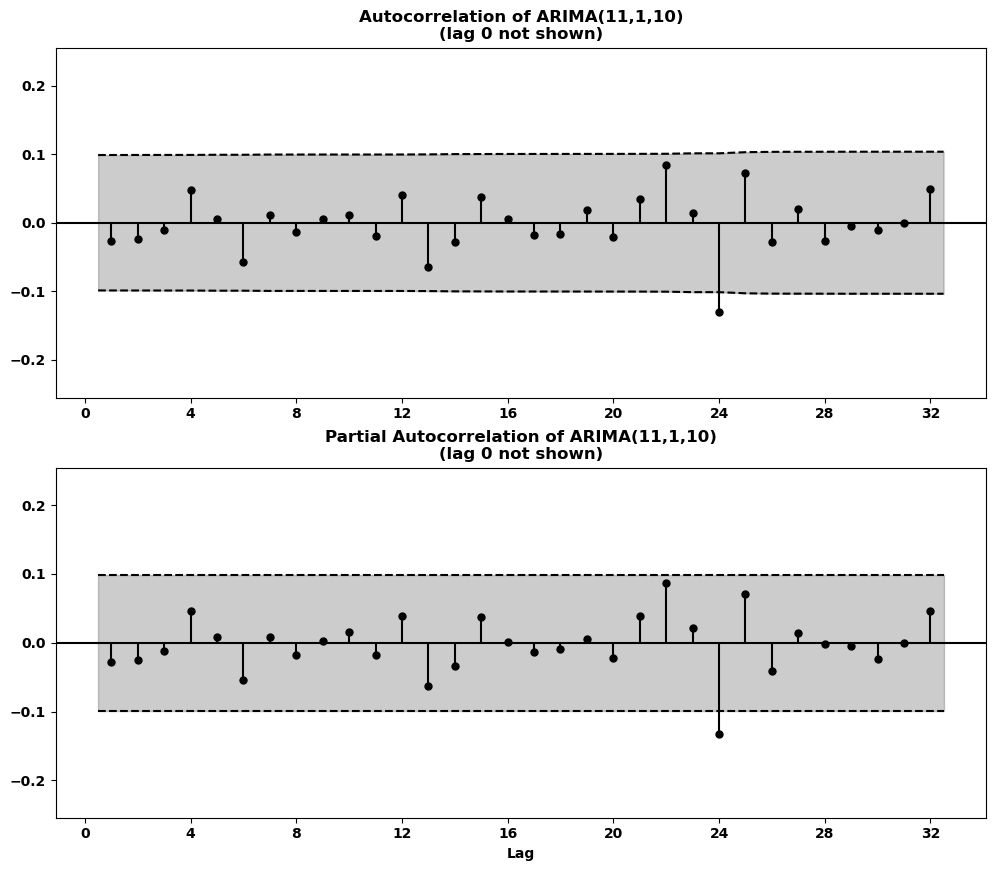
# Results

## Traditional ARIMA Models

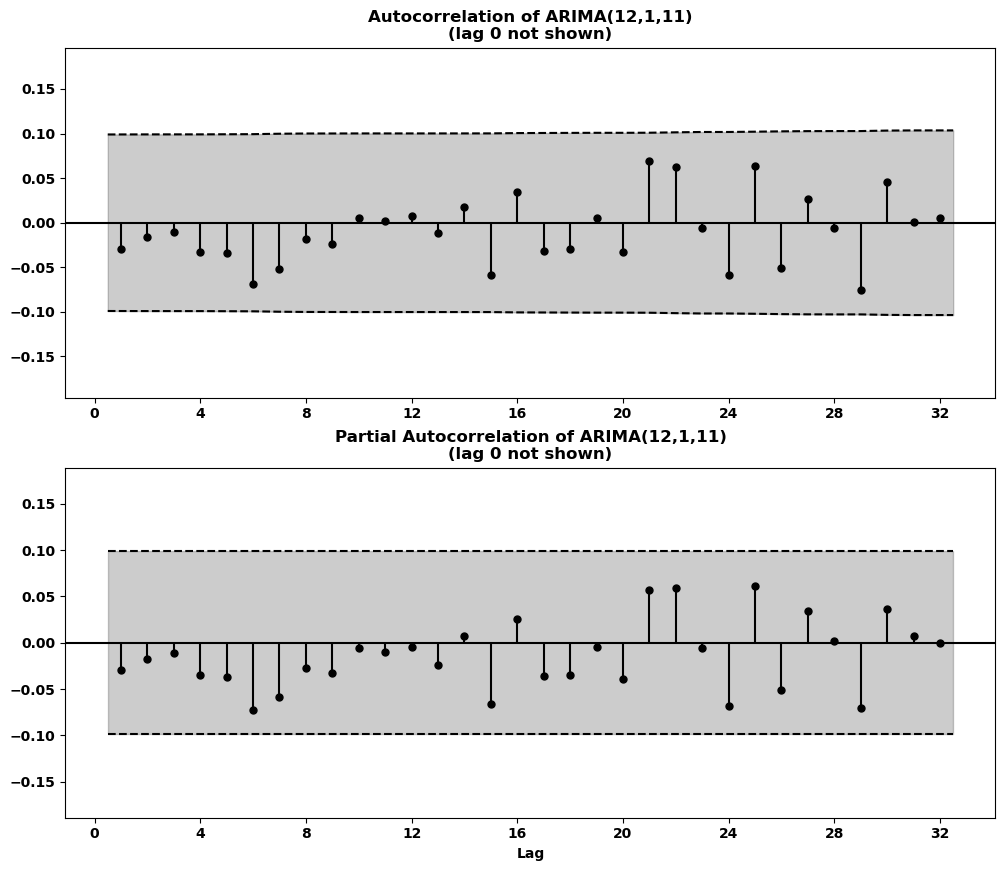
We present some traditional forecasting models with which to compare our wavelet-based forecasting. We employed a programmatic loop that iteratively finds the model specification that gives the lowest root mean square error in levels to take the guesswork out of the equation. We also employed a similar loop that finds the model with the lowest information criteria (IC); however, the model with lowest IC did not exhibit relatively low root mean square error, we decide to omit that model and instead focus on the top two performers from an RMSE accuracy standpoint. These loops consider both pure ARIMA models as well as models with monthly dummy variables, which we think could capture some seasonality within the oil price series.

For the ARIMA series, which will act as the baseline, we find an ARIMA (12,1,11) with no monthly dummies, to have the lowest root mean square error of 13.3384. Its BIC, on the other hand, is 2272.3000. Next, we employ the Diebold-Mariano test to see if this model generates statistically different forecasts than the next-best model, which is an ARIMA(11,1,10) with monthly dummies. The test fails to reject the null at the 5% level, indicating that these forecasts are not statistically different. To discern which model should be the model put forth for comparison, we turn to residual analysis.

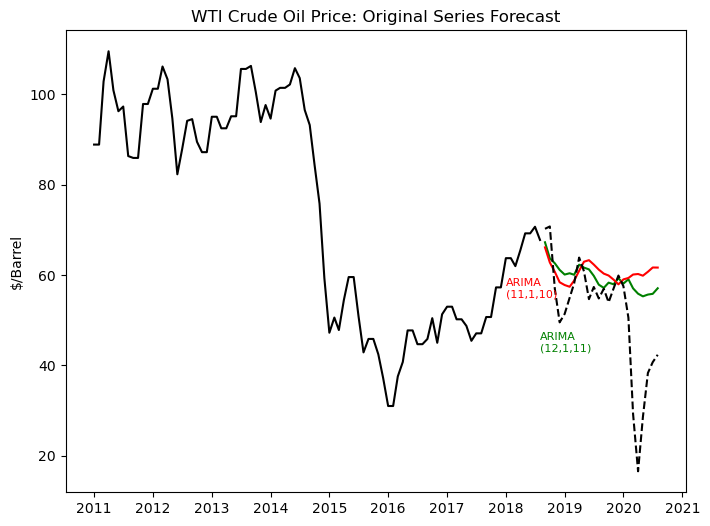
Conducting a correlogram analysis of the ARIMA(11,1,10) model shows a significant spike beyond the 95% bands for both the sample autocorrelation and the sample partial autocorrelation, hinting at some bi-yearly seasonality still unaccounted by the model.



By contrast, the ARIMA(12,1,11) model does not share this similar problem, perhaps due to its twelve autoregressive lags



In conclusion, the parsimony principle would lead us to select the ARIMA(11,1,10) model, however residual analysis shows that the larger ARIMA(12,1,11) model better accounts for all seasonality in its residuals. This is reflected also in the latter model’s lower BIC. Thus, we put forth both models for comparison. The forecasts from these two models plotted below:



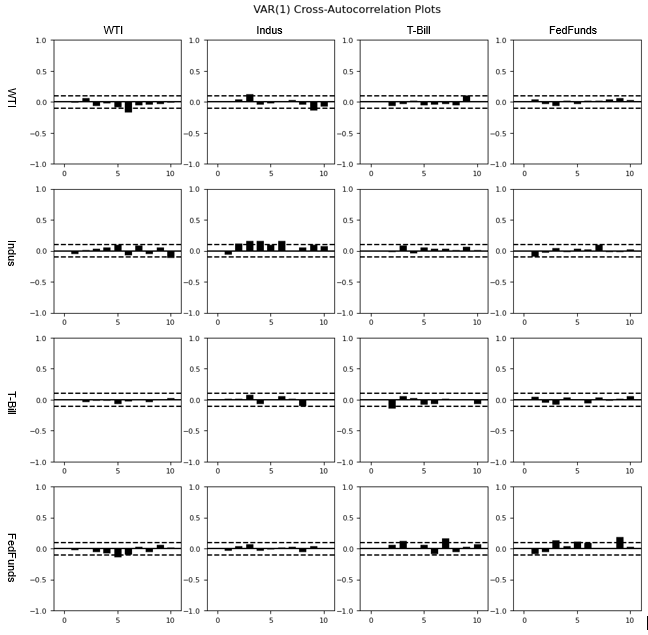
## Traditional Vector Autoregressive Models

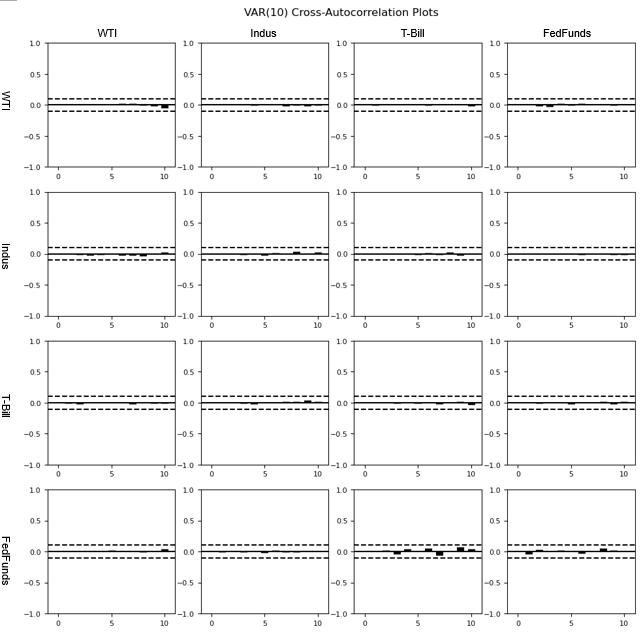
Given how oil prices are instrumental for and heavily intertwined with the pricing of other critical financial prices, we also provide here a vector autoregressive (VAR) model for comparison. We decide on three variables with which to estimate and model WTI oil prices: Industrial Production Index, 10-Year Treasury Bonds, and the Federal Funds Effective Rate.

As with the methodology above, we employed a similar programmatic loop to find the specification with the lowest mean RMSE among all four series to find a VAR(10) model, noting that the selected specification did not change even if we only prioritized the RMSE of the dependent variable of WTI oil prices. We also employed a loop to find the model with the lowest information criteria. The model with the lowest AIC is a VAR(6), and the model with the lowest BIC is a VAR(1). As the information criteria chose different models, we choose the VAR(1) model with the lower BIC instead of the lower AIC at the advice of Diebold (2007). We compare this model against the VAR(10) model which gives the lowest mean RMSE.

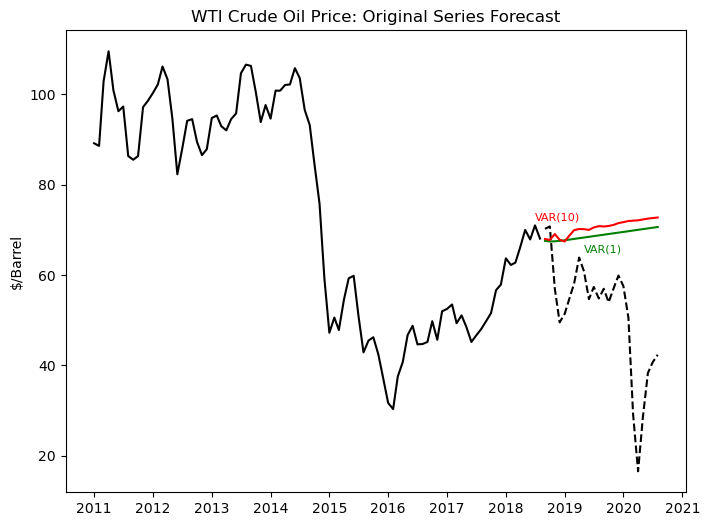
Testing the forecasts against the validation set, the VAR(1) model gives a mean RMSE of 70.3342 while the VAR(10) model gives us a mean RMSE of 71.9813. We were unable to reject the null of statistical identical forecast accuracy in the Diebold-Mariano test, but before opting for the more parsimonious model as the best model, we first turn to residual analysis of these models. Considering the VAR pairwise cross-correlation plots. Overall, there is still seasonality left unaccounted for in this parsimonious model. The top left-bottom right diagonal represents the autocorrelation of the series itself. All of the variables exhibit unaccounted autocorrelation in its own residuals. By contrast, most of the plots which delineate the correlation between any particular series and the lags of any other series do not have spikes breach the 95% Bartlett threshold. The only exceptions are in the relationship between the Federal funds effective rate and the lags of WTI oil prices and the 10-Year Treasury Bond rates (bottom row, first column; and bottom row, third column; respectively).

On the other hand, the cross-autocorrelation plots of the VAR(10) model does not exhibit any unaccounted seasonality in the residuals of any series onto itself or any other of the three series jointly estimated. Thus, we will choose the VAR(10) option due to its better fit, even if we sacrifice a little accuracy in the process.



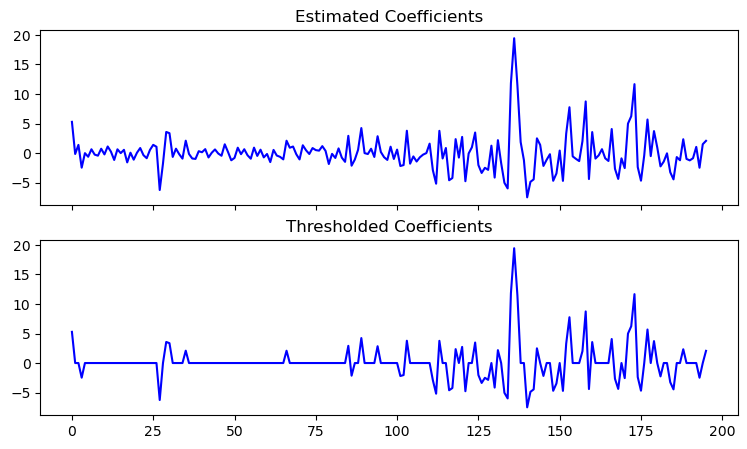


Below are the forecast plots between the VAR(1) and the VAR(10). We can see that the forecasts perform extremely poorly for both models, which were unable to find the downward progression of the oil prices whatsoever. This is because the other series do not exhibit the same price fluctuations as our dependent variable, and as such may have “washed out” the effect of the dependent variable on the VAR model overall.



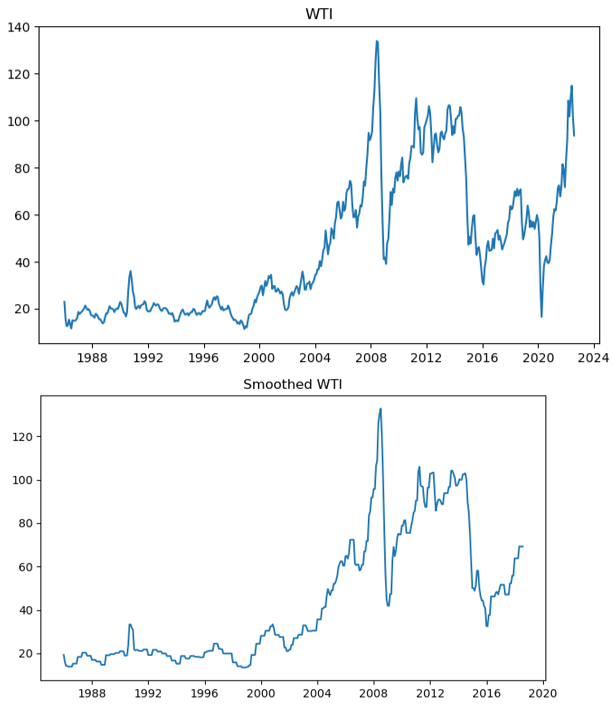
## Smoothed Series ARIMA Model

The first wavelet-based forecasting method that we try is forecasting the smoothed oil prices series; see Figure 4 for a notional illustration. A common application for wavelet decomposition is denoising or smoothing signals. For example, many image denoising algorithms leverage wavelet methods. We employed a standard smoothing technique on the training set of our WTI oil prices data. We first decompose the signal down two levels, then we take the two series of wavelet coefficients that are produced and apply a threshold. All coefficients withs values less than the threshold are set to zero. The idea is that small variations, which get represented by small coefficients are not actually meaningful signal, or stated more specifically in our context for forecasting, these small variations are not useful for predicting future values of the series. Only large variations, which are not thresholded, are the most useful for forecasting. An example of the coefficient thresholding is pictured below:



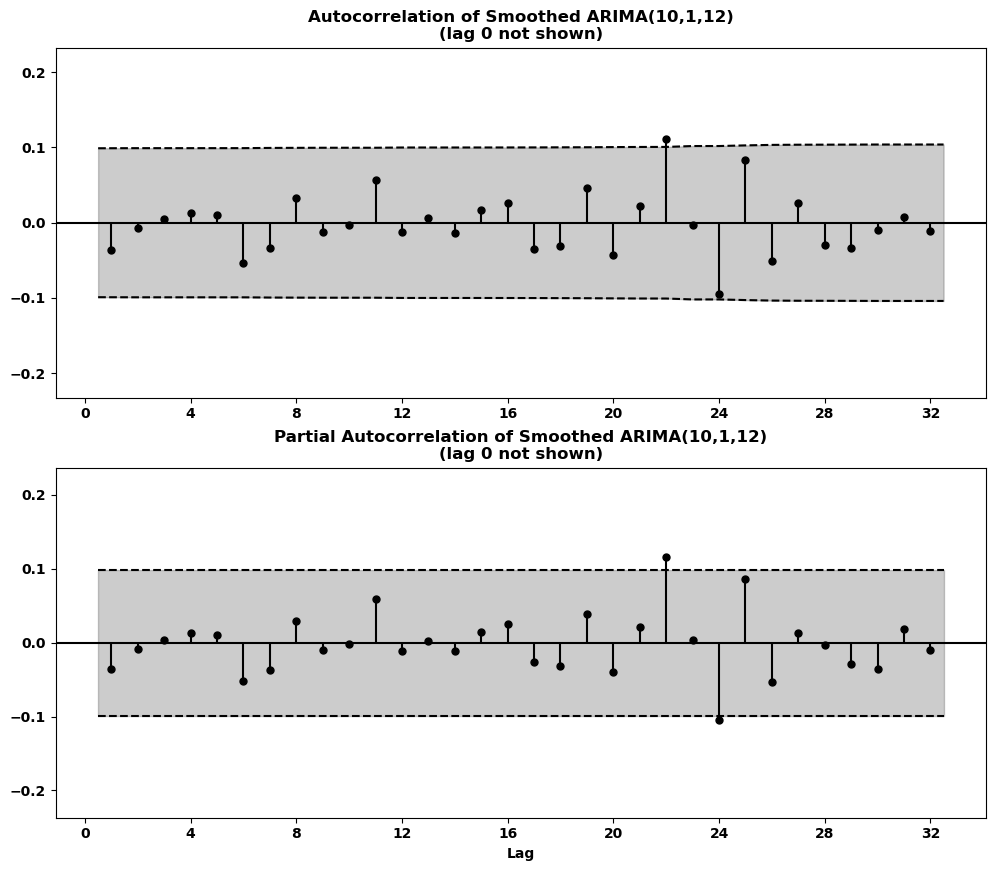
After zeroing-out most of the coefficients, then we apply the inverse wavelet transform to obtain a smoothed version of the original series. Note that *only* the training set needs to be smoothed in this way. When using the smoothed training set to predict future values of the series, we evaluate these forecasts against the validation set of the original, non-smooth series. In the denoising algorithm, we use a Haar wavelet and choose the threshold to be , which are reasonable defaults for denoising applications. More experiments could be done to determine the best threshold parameter, or to study the interaction between the choice of threshold, choice of wavelet function, and the selected ARIMA model orders, but that is beyond the scope of this paper.

For reference, here is a comparison between the original series and the smoothed series. Visually, the smoothing created “flat” peaks that once were jagged. The smoothing also loses the dip towards the very end of the smoothed series.

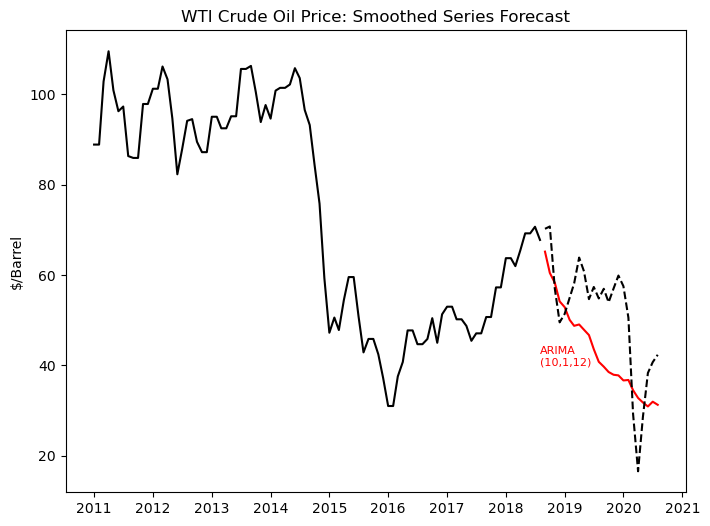


This smoothed series is then fit with an ARIMA model, in the same manner that the original oil prices series can be fit with ARIMA. We found through programmatic loops that an ARIMA(10,1,12) without monthly dummies provided the lowest RMSE when tested against validation data. It had an RMSE of 10.640 and a BIC of 2269.775. Using the Diebold-Mariano test, we find that the forecast generated by this model was found to be statistically different to that of the next-best model, an ARIMA(12,1,11) with monthly dummies.

Recall that for the original series, we selected ARIMA(11,1,10) with monthly dummies as the best forecast model. Here, we select a model with similar ARIMA orders--ARIMA(10,1,12) versus ARIMA(11,1,10)--although, interestingly the best model with the smoothed series does not include monthly dummies. Correlogram analysis indicates very little remaining seasonality unaccounted for in the residuals of the ARIMA(10,1,12) model. Perhaps the only hint of seasonality lies in the spikes which are very close to the threshold at rather high lags, so we are willing to tolerate those. Below are the correlograms:



Below is the forecast plot of the ARIMA(10,1,12) series. We can see that the model finds the downward progression correctly, although it is unable to mimic the dynamics more closely. It also fails to predict the large spike upward around mid-2019.

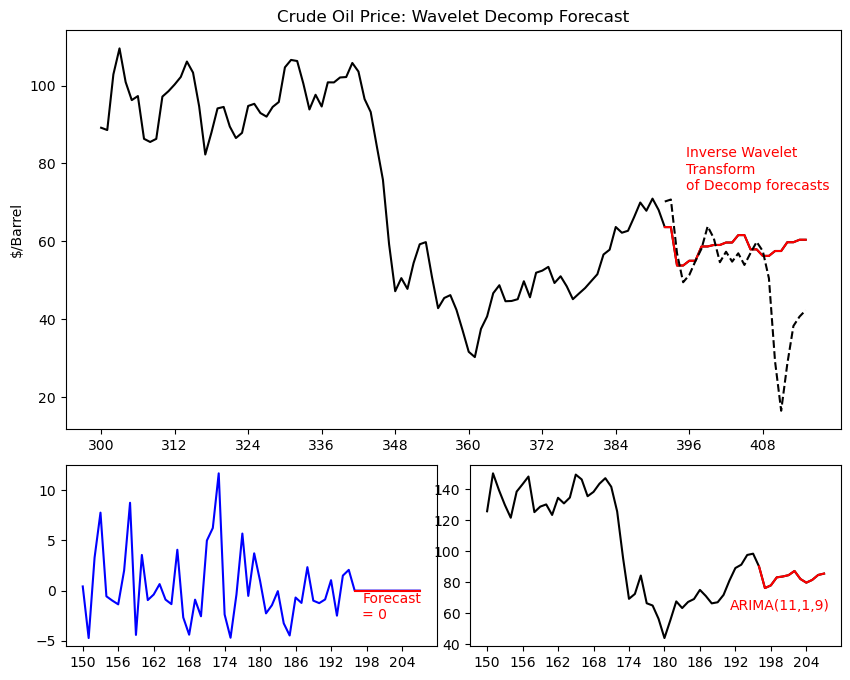


## Wavelet Decomposition ARIMA Models

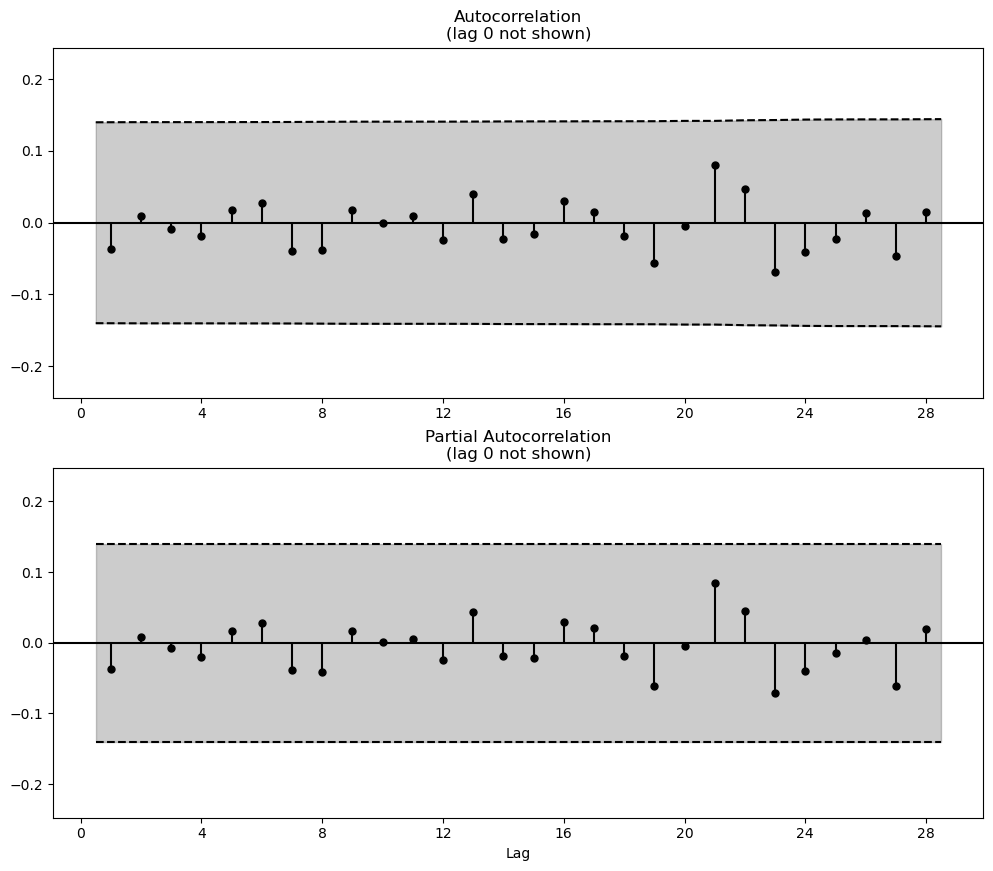
The second wavelet-based forecasting method that we try is to forecast the wavelet multiresolution decomposition components; see Figure 5 for a notional illustration. For this approach, we decompose the original series by one level, using a Haar wavelet. This produces two series that now need to be forecasted in order to reconstruct a forecast for the original oil prices. After some initial exploration of model fitting, we noticed that the forecasts of the coefficient values themselves are essentially zero. This actually makes perfects sense, because the coefficients represent small deviations from the true signal, i.e. noise! Of course, the best forecast of a series of noise with mean zero would be zero!

With this realization, we can focus only on searching for the best model for the level one series decomposition. We simply set the “forecasted” coefficients to zero. Note that only the training set is transformed with the wavelet decomposition. We use the reconstructed forecast on the original series to evaluate versus the validation set. To obtain 24-step forecast on the original series, we need to forecast only 12 values of the level one decomposition. As with the other experiments, we test models with and without monthly dummies, although in this case, we construct seasonal dummies with a period of 6 to be exogenous variables in forecasting the level one decomposition.

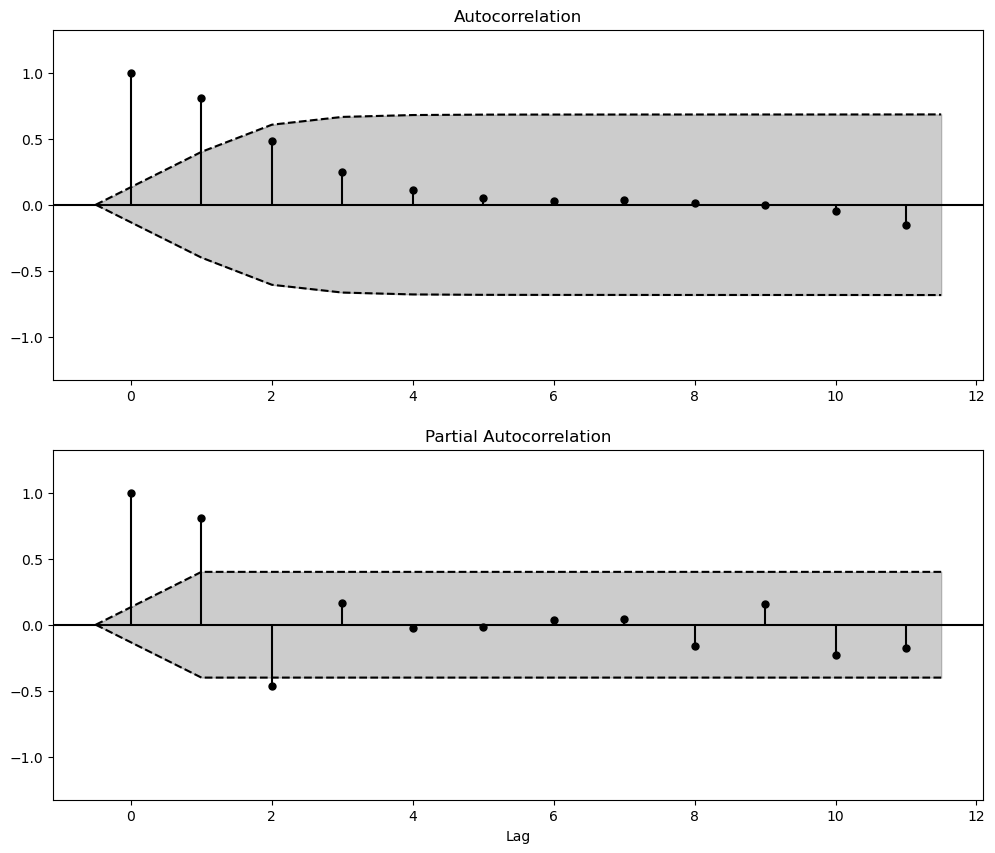
The model with lowest RMSE is ARIMA(11,1,9) with seasonal dummies; the RMSE is 14.389. The BIC for this model is 1471.69, although note that this IC is not comparable to the IC calculated on the original series.



Below, we see the ACF and PACF plots for the residuals of the fit of the ARIMA(11,1,9) model to the level-one decomposition (the fit to the series in the lower-right corner of the plot above). The residuals do exhibit any lingering correlation with lagged values.



In the next plot, we show the ACF and PACF plot of the forecast residuals. We can only show a small number of lags because the forecast horizon is only 24 steps. We see significant autocorrelation at lag 1 and partial correlation each at lag 1 and 2, which indicates that this model is not capturing all the possible forecastable information in the series.



## Model Comparison

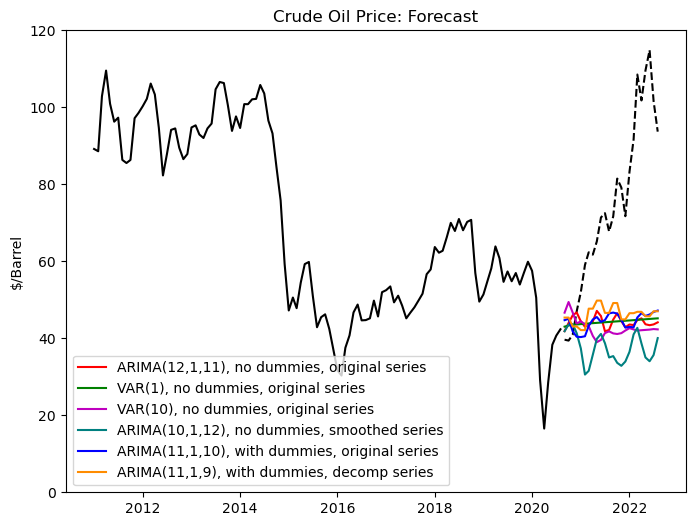
Table 1: Summary of model performance on validation set

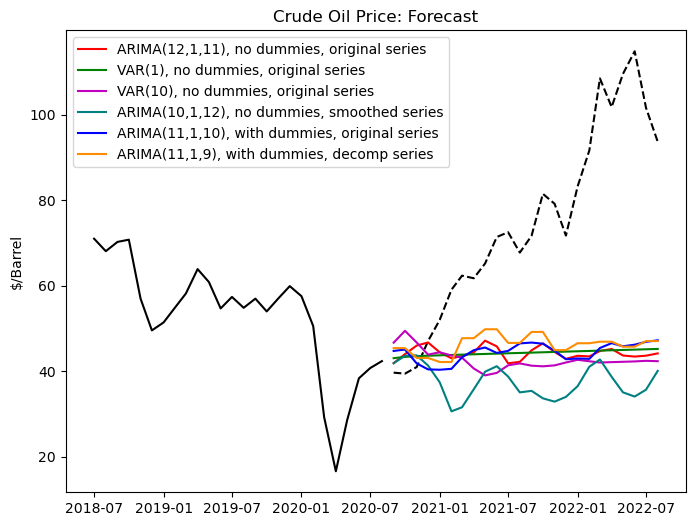
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Technique | Selected Model | Monthly Binary Variable? | RMSE | BIC |
| Baseline | ARIMA(11,1,10) | yes | 13.349 | 2314.57 |
| Baseline | ARIMA(12,1,11) | no | 13.338 | 2272.30 |
| Baseline | VAR(1) | no | 71.981 | -4.725 |
| Baseline | VAR(10) | no | 70.334 | -4.032 |
| Smoothed series | ARIMA(10,1,12) | no | 10.640 | 2269.775 |
| Wavelet decomposition | ARIMA(11,1,9) | yes | 14.389 | 1471.69 |

Having used the validation set to choose the best candidate models for forecasting oil prices, we now move on to testing these candidates against the final hold-out set. We re-train the models from Table 1 on the training set now concatenated with the validation set. We do this only to update the estimated model coefficients for the models listed in Table 1; we do not perform any additional search for candidate models. The results of forecasting versus the hold-out test set are presented in Table 2, and plots of all the candidate forecasts are presented below as well.

Table 2: Summary of model performance on testing set

|  |  |  |  |
| --- | --- | --- | --- |
| Technique | Selected Model | Monthly Binary Variable? | RMSE |
| Baseline | ARIMA(11,1,10) | yes | 36.754 |
| Baseline | ARIMA(12,1,11) | no | 37.645 |
| Baseline | VAR(1) | no | 37.192 |
| Baseline | VAR(10) | no | 39.512 |
| Smoothed series | ARIMA(10,1,12) | no | 43.790 |
| Wavelet decomposition | ARIMA(11,1,9) | yes | 35.563 |





Finally, we ran a series of pairwise Diebold-Marino tests on the testing set forecasts. The RMSE for models ARIMA(11,1,10), ARIMA(12,1,11), and VAR(1) are statistically indistinguishable from each other. The RMSE for models VAR(10), ARIMA(10,1,2) on smoothed series, and ARIMA(11,1,9) on the decomposition series are each statistically different from all the other models.

# Conclusions

By happenstance, our ultimate testing set exhibits a sudden sharp upward trend, which is hard to forecast with any model. On one hand, we have a testing scenario that is reflective of a very realistic forecasting scenario, but on the other hand, the difficulty of the testing set makes it hard to determine if the added complexity of the wavelet-based methods may actually improve forecasts. Nevertheless, we can make some final comments on our results.

Forecasting the smoothed series performs the worst on the test set, despite the fact that it was the best performer on the validation set. It appears that this approach can *over-smooth* the training data. The most complex method, forecasting on the wavelet decomposition, performed the best on the test set, notably better than standard baseline ARIMA models. This is consistent with the conclusions of Schlüter and Deuschle (2010), where they found for a 5-trading-day forecast for daily oil prices, the forecast on the wavelet multiresolution decomposition had lower RMSE than baseline ARIMA models. Qualitatively, we see in the plots comparing all six candidate forecasting models, the decomposition method (orange curve) captures the upward trend of the oil prices better than any other model, and particularly so for the first half of the forecast horizon. Therefore, we recommend that if researchers or practitioners are not deterred by the complexity of the method, they can add wavelet multiresolution decomposition as an another tool in their forecasting toolbox.

## Future Work

Our conclusions for the merits of wavelet forecasting were made based on only one example economic series. Now that we have engineered the testing pipeline, a natural next step would be to programmatically test many more series. This would help determine whether wavelet methods are likely to provide good forecasts in a general case. Further, if wavelet forecast methods work for some series and not for others, can we determine what qualities of those series make them attractive for wavelet methods?

A large-scale study could also assess the forecast performance on a rolling forecast horizon. Are there universal change-points where wavelet methods work well for certain forecast horizons, or is it series dependent? Ideally, a large-scale study could also explore forecasting various time series frequencies. Do wavelet methods work better for low or high-frequency series?

Finally, further work could be done to improve the forecasting of the wavelet multiresolution components. This work aligns with that of Schlüter and Deuschle (2010) which defaulted to applying ARIMA type models to all series. However, for forecasting series of wavelet coefficients (particularly for coefficients that have been thresholded), a method such as that found in Croston (1972) that works well for series with many zero values, could improve the performance of the wavelet-based forecasting methods.

# References

Croston, JD. Forecasting and stock control for intermittent demands. *Operational Research Quarterly*, 23(3), pp. 289–303, 1972.

Diebold, F.X. *Elements of forecasting*. United Kingdom: Cengage Learning. 2007.

“Federal Funds Effective Rate.” Federal Reserve Economic Data. May 1, 2023. https://fred.stlouisfed.org/series/FEDFUNDS

“Industrial Production and Capacity Utilization - G.17”. Federal Reserve Economic Data. April 14, 2023. https://www.federalreserve.gov/releases/g17/About.htm

Lee et al., “PyWavelets: A Python package for wavelet analysis.” *Journal of Open Source Software*, 4(36), pp. 12-37, 2019. doi: 10.21105/joss.01237.

“Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis.” Federal Reserve Economic Data. May 1, 2023. https://fred.stlouisfed.org/series/GS10

Mallat, S.G. "A theory for multiresolution signal decomposition: the wavelet representation," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674-693, July 1989. doi: 10.1109/34.192463.

Python Software Foundation. Python Language Reference, version 3.8.12, <http://www.python.org.>

Schlüter, S. & Deuschle, C. “Using wavelets for time series forecasting: Does it pay off?” IWQW Discussion Papers, No. 04/2010, Friedrich Alexander-Universität Erlangen-Nürnberg, Institut für Wirtschaftspolitik und Quantitative Wirtschaftsforschung (IWQW), Nürnberg, 2010. http://hdl.handle.net/10419/36698.

Seabold & Perktold, “Statsmodels: Econometric and statistical modeling with python.” *Proceedings of the 9th Python in Science Conference*, 2010. <http://conference.scipy.org/proceedings/scipy2010/pdfs/seabold.pdf>

U.S. Energy Information Administration, Crude Oil Prices: West Texas Intermediate (WTI) - Cushing, Oklahoma [MCOILWTICO], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/MCOILWTICO, April 9, 2023.