December 11, 2016

Zsófia Farsang - András Kárpáti

machine learning, 2016 Fall – Milán Badics

Rajk László College for Advanced Studies

Forecasting oil price volatility

data is the new oil, oil data is the new oil oil

# Abstract

Oil prices determine a variety of economic and financial factors, and their historical volatility and proneness to shocks makes forecasting them essential for analysts and economists globally. Oil prices are influenced by the supply-demand equilibrium, exogenous effects such as war, and endogenous effects. Research of the topic has dwelled into all three of the categories, and utilized different econometric and computational intelligence models to predict oil price levels and volatility. Our paper aims to build on previous research, and forecast oil price volatility by using a combination of linear and non-linear models. We forecast with the autoregressive conditional heteroscedasticity (ARCH) model, and artificial neural networks.

Our analysis includes Google Trends data of select keywords as inputs. After our analysis, we find that the neural network is a better predictor for the volatility for the next 4 weeks than any other model specified. Regressing neural network to predict the residuals of the ARCH models does not significantly improve RMSE, with the Google Search data similarly insignificant in improving the RMSE. Overfitting problems are possible present in the model, giving us inconclusive results. Overall, our model tends to underestimate big shocks, and overestimate the volatility of smoother periods.

Table of Contents

[Abstract 2](#_Toc469242065)

[1 Introduction 1](#_Toc469242066)

[2 Oil price volatility and previous research 2](#_Toc469242067)

[2.1 Literature review 4](#_Toc469242068)

[2.1.1 Other factors behind oil prices 5](#_Toc469242069)

[2.1.2 Methodologies used in forecasting oil prices 5](#_Toc469242070)

[3 Methodology 8](#_Toc469242071)

[3.1 Our contribution to the topic 8](#_Toc469242072)

[3.2 Hypothesis 8](#_Toc469242073)

[3.3 Database construction 8](#_Toc469242074)

[3.4 Regression methods 11](#_Toc469242075)

[3.4.1 ARCH 11](#_Toc469242076)

[3.4.2 Neural Network 11](#_Toc469242077)

[3.4.3 Neural network combined with ARCH 11](#_Toc469242078)

[4 Results 11](#_Toc469242079)

[4.1 Fit to data 11](#_Toc469242080)

[4.2 Predictive power 15](#_Toc469242081)

[4.3 Further research 16](#_Toc469242082)

[5 Conclusion 17](#_Toc469242083)

[6 Table of Figures 19](#_Toc469242084)

[7 References 20](#_Toc469242085)

# Introduction

Oil is the backbone of the modern economy, with countless industries and companies depending on it as a production input. Oil prices affect growth, inflation and manufacturing capabilities, and their volatility is a risk every modern economy must count with. Oil is a non-renewable resource distributed unevenly around the globe, but demanded and traded almost universally, with developing economies only recently starting to exhibit the higher needs for crude oil and further driving up demand. As researched by Hamilton (2011, pp. 19), a subset of newly industrialized economies were responsible for 17% of global oil consumption in 1998, and accounted for 69% of the increase of oil consumption between 1998 and 2011. Due to its limited and nonrenewable nature, oil supply in the long-term has serious barriers. Most experts feel that this increasing gap between demand and supply will result in increasing prices and due to the unpredictability of the aforementioned factors, higher volatility.

Oil price shocks have influenced our economic growth several times in the past, we need only think of the 1973 and 1979 oil price shocks, or the rising prices and volatility of the mid-2000s. Academia and business both feel the urge to find reliable and efficient ways to forecast oil prices, and thus prepare and mitigate the risks its volatility means to our economies and businesses. Various methodologies have been used in the past to predict prices, and the aim of our paper is to investigate some of these approaches, while building on them with our own model. Our study focuses on oil price volatility, based on weekly WTI crude oil spot prices. The basic model predicts the prices by using the ARCH model, followed by the utilization of a neural network. Our hypothesis in this paper is that the neural network outperforms the ARCH forecast, and the addition of Google search histories of relevant keywords improves the accuracy of the forecast.

Our paper is organized as follows: Chapter 2 introduces historical trends in crude oil prices and their volatility, followed by a short summary of selected research into the field of oil price and volatility forecasting. Chapter 3 provides an overview of the methodologies used in the paper, including the ARCH, GARCH and neural network models. In Chapter 4 we show the results of our research, and its fit with the previously introduced hypothesis. Chapter 5 concludes our paper.

# Oil price volatility and previous research

In order to correctly model and forecast oil price shocks and price volatility, we have to grasp the factors behind their dynamics. Oil prices are greatly determined by supply and demand, including precautionary demand - demand determined by our expectations of future supplies. (Hamilton, 2009). Forecasting oil price levels is important for the actors of an economy to plan their actions and their output, but volatility cannot be overlooked either. Oil price volatility plays an enormous role as an input for macroeconomic models, option pricing formulas and various derivatives. Volatility dynamics have a relationship with economic growth, inflation, energy futures contracts and other assets. Modelling this dynamic and forecasting the volatility is thus a powerful tool that can help economic actors mitigate their risks. (Kang, Kang and Yoon, 2009).

This volatility is clearly present on Figure 1., which shows the weekly average spot FOB price of Cushing, OK WTI per barrel between 1986 and 2016. From Figure 1 we can observe the steady rise of crude oil prices until 2008, where another oil shock sent the price tumbling down from $145 per barrel. The spot price is showing high volatility from the beginning of the observed period, which gets progressively worse from the year 2000 onwards.

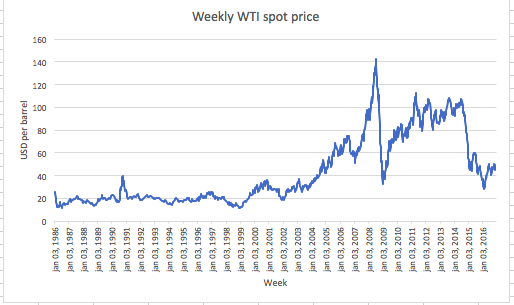


Figure 1 - WTI spot prices between 1986 and 2006, in USD per barrel  
Source: U.S. Energy Information Administration, 2016, n.pag.

Previous work in this field has identified three main factors that contribute to the volatility of crude oil prices: The first factor is increasing demand and shortages in oil supply, which can be attributed to economic growth of oil-importing countries; and the production behavior of oil-exporting countries.  The second factor is exogenous effects, such as political instability, wars or catastrophes in the oil-producing regions. The third factor is attributed to endogenous reasons, e.g. speculators. A vast amount of studies into oil price forecasting have focused on the relationship between demand, supply, inventories and oil price, their results hinting at other factors as possible predictors of oil price. (Pan, Haidar and Kulkarni, 2009). This categorization was supported by Wang and Wang (2011), who used neural network and kernel density estimation.

Correlation of events that exhibit the nature of the second factor - political upheaval, instability, war, or catastrophes - can be observed by looking at historical data of oil prices, and the timeline of major events concerning the oil-producing countries. Table 1. summarizes the major oil price shocks after 1973, as put forward by Hamilton (2011). We can observe that most of them had political unrest in oil producing countries (the Middle East and Venezuela) behind them, with the minority of cases where demand and supply were acting mostly on their own.

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Event | Production effect | Price effect |
| October 1973 | Yom Kippur war - Egypt and Syria attack Israel  OAPEC members put an embargo on countries supporting Israel  Strong demand | Arab countries production shortfall 7.5% of global output | 51% price increase between Nov-Feb |
| Early 1979 | Iranian revolution | Production shortfall of Iran 7% of global output | 57% price increase between May 79 - Jan 80 |
| September 1980 | Iran-Iraq war | Production shortfall of Iran and Iraq 6% of world production | 45% price increase between Nov 80 - Feb 81 |
| 1990 | Gulf War I - Iraq invades Kuwait | Iraqi and Kuwait production collapses  Saudi excess capacity restores world production | 93% price increase between Aug-Oct |
| 1999-2000 | Strong demand, global economic downturn | - | 38% price increase between Dec 99 - Nov 00 |
| 2002-2003 | Venezuela unrest caused by strikes, US attack on Iraq | Production fall of Venezueal and Iraq counterbalanced by growing global output | 28% price increase between Nov 02 - Mar 03 |
| 2007-2008 | Strong demand  Stagnant supply  - no growth after 2005, continued unrest in Iraq and Nigeria, fields reaching maturity | - | 145% price increase between Feb 07 - Jun 08 |

Table 1 - Major oil price shocks between 1973 and 2008, and their effect on production and prices. [[1]](#footnote-1)

Source: own table, based on Hamilton, 2011, pp. 33. For production level data, Hamilton, 2011, pp. 14; 16; 17 respectively.

## Literature review

As shown in the previous sections, forecasting oil prices accurately has been in the focus of a myriad researchers and analysts in the last few decades. In order to better understand the body of work written on the topic, first we have to provide a quick overview of the assumptions on what factors and indicators have predictive capabilities for oil price; then we will briefly summarize the different econometric and computational intelligence models used in forecasting oil prices.

### Other factors behind oil prices

Oil price fluctuation - commonplace events even in their extremes, as introduced in the first part of the paper - can have a dramatic effect on both supplying and consuming countries of oil. Research has been done into the relationship between several economic factors and oil future prices. This section aims to introduce an extended view on the basic framework of supply-demand, exogenous and endogenous effects introduced in the previous section.

Lizardo and Mollick (2010) have found an inverse relationship between the oil price and the value of the U.S. dollar, while Bénassy-Quéré, Mignon and Penot (2007) found that the increase of the oil price leads to the appreciation of the U.S. dollar. Crude oil price volatility have been found to have an impact of 6% on stock market returns, as investigated by Park and Ratti (2008, cited by Abraham and Gabralla, 2013, pp.731) on returns of the U.S. and 13 European countries.

Based on previous research, Abraham and Gabralla (2013) have identified several other factors that predict the decline or rise in oil prices: gold, international oil supply, international oil demand, political factors, natural disasters, OPEC policy and future contracts. Cold weather, OPEC decisions about raising prices, war or revolution in the oil producing countries, the price of gold and non-OPEC decisions about cutting production lead to an increase in the oil price. Decisions about raising production results in the devaluation of the oil price. Speculators, exchange rate volatility and the future contract market enhance the volatility of oil prices.

### Methodologies used in forecasting oil prices

Approaches to forecasting have been dependent on the available technology and methodologies, which have improved vastly in the last few decades. In this section, we provide a quick overview of the different models used by researchers in the past and their success, in order to develop a model for our own forecast.

Previous studies focused on oil market prediction can be grouped into 1) models that use future prices as predictors of spot prices; 2) econometrical models that attempt to explain and predict, and 3) computational intelligence models aimed at prediction. (Pan, Haidar and Kulkarni, 2009). Another approach is sorting models into quantitative and qualitative methods, as done by Xie et al. (2006). Abraham and Gabralla (2013) divide the existing research into the following three categories: 1) correlation between oil price and economic variables, 2) characteristics of main factors, 3) oil price volatility and forecasting models. There are significant overlaps between these three groupings, thus we will only elaborate on the first two approaches (Pan and Xie, as introduced before), as they show quite clearly the distinction between the various models.  First, we briefly describe the difference between quantitative and qualitative models, than separate previous research into the linear and non-linear categories, as seen in Pan, Haidar and Kulkarni (2009).

We can sort models into the categories of qualitative and quantitative. Qualitative models rely heavily on human judgement, and have problems that lead to their performance being inferior to quantitative models. Quantitative methods either rely on the supply and demand equilibrium (structure methods), or use time series data to forecast oil prices (data-driven methods) (Abraham - Gabralla, 2013).

#### Linear models

Research modelling spot prices with futures prices utilizes the fact that information affects future prices quicker than spot prices, as trading has lower transactional costs and future markets tend to be more liquid. Despite the body of work in the subject - mainly econometric models -, there is no agreement among researchers on the predictive relationship between future and spot prices. (Pan, Haidar and Kulkarni, 2009).

Econometric models have developed in analyzing and predicting oil prices in the last decades, especially with the invention of GARCH and ARIMA models. The econometric approach’s fault lies in its assumption of linearity, which causes the models to fail in the case of non-linearity. (Pan, Haidar and Kulkarni, 2009). These linear time-series - such as ARCH, GARCH or ARMA models - are often used in predicting oil prices. The difficulty with using these methods is the underlying assumption of a linear connection between the crude oil price and the variable(s) used in predicting it. These approaches have been found superior to a random walk model, as seen in the case of GARCH by Morana (2012), and as seen for vector error correction by Coppola (2008). In reality, these models on their own prove impractical, as crude oil price and its volatility is a much more chaotic phenomenon, which contains non-linearity. (Abraham and Gabralla, 2013).

GARCH models and its variants were used by others to model and forecast crude oil price volatility. However, standard GARCH models on their own are not capable of capturing volatility persistence. Kang, Kang and Yoon (2009) have analyzed the volatility of WTI, Brend and and Dubai crude oil prices, and found that modified GARCH models (FIGARCH and CGARCH, which both consider long memory) forecast with better results than the standard GARCH model. Expanding on their model, Wei, Wang and Huang (2010) use a greater number of linear and non-linear models. Although the paper does not reach the same result of some non-linear GARCH models outperforming the others, the authors also conclude that non-linear GARCH models that can capture long memory, are more accurate than their linear counterparts. Their results are even more emphasized on longer forecasting horizons (5-20 days).

#### Non-linear models

As for computational intelligence models, artificial neural networks (ANN), support vector machines (SVM), genetic algorithms (GA) and case-based reasoning (CR) have all been utilized in the field. (Pan, Haidar and Kulkarni, 2009). Moshiri and Foroutan (2006) have tested the linearity of crude oil futures prices, their results showing the futures prices time series to be non-linear. Their study compared linear (ARIMA and GARCH) and non-linear (ANN) models in their ability to predict futures prices, where ANN has proven to be statistically significant, and superior to the linear models. Xie at al. (2006) found that ANN and SVM both outperformed ARIMA models on out-of-sample data when predicting monthly crude oil prices.

Yu et.al.’s (2008) used an EDM-based neural network ensemble model has shown better results than comparable options. Pan, Haidar and Kulkarni (2009, pp. 177) use artificial neural networks to forecast 1,2 and 3-day oil prices with an out-of-sample success rate of 80%, 70% and 61%. Their model uses historical oil prices, oil futures prices and intermarket information.

Despite successes, neural network models are prone to overfitting, poor generalization performance and do poorly in determining appropriate network structures. (Abraham and Gabralla, 2013) A criticism regarding the usage of ANN lies in the foundations of the methodology. ANN generates better networks from more available data points in general, but this is doubtful in the case of financial or economic time series. Financial and economic data become irrelevant with time - as the information becomes old, it has less influence on the behavior on the current actors and the current market. Using outdated data as a training set for ANN can impact its ability to successfully generalize. (Pan, Haidar and Kulkarni, 2009).

# Methodology

In the following chapter, we propose the model of our paper, which uses ARCH, ANN and Google Trends data to predict oil price volatility. Using these models we test our hypotheses about the predictive powers of each specification. The third section describes the database used in forecasting, followed by the summary of the methodologies used in the paper in section 4.

## Our contribution to the topic

As we saw in the literature review, forecasting volatility of financial time series is widely researched both in econometrics and machine learning. Our contribution to the literature is that we try to forecast a specific time series – oil price – and use a proprietary dataset containing several important Google Search keywords.

Our idea is that Google Search can be a good predictor of short term volatility, since usually people search for terms that interests them, and the intensity of this interest is quantified in the number of searches they do. As discussed above, higher volatility periods also occur when exogenous effects (as war and catastrophes) happen, which usually make headlines and generate more searches. The same stands for negotiations and decisions of oil supply by e.g. OPEC concerning serious volumes.

## Hypothesis

Our hypothesis is threefold. Our main hypothesis is that a neural network can capture the nonlinearity of oil price volatility, therefore the neural network prediction should be more accurate than the ARCH models. Second, we hypothesize that including Google Trends data of select key words will improve the forecast. Third, we think that estimating the residuals of the GARCH estimation will improve the forecast.

## Database construction

We downloaded weekly WTI (West Texas Intermediate) USD / barrel prices from the US. Energy Administration Website (EIA) containing 262 observations for the past 5 years as our main variable (U.S. Energy Information Administration, 2016).

The explanatory variables based on the literature review were:

·    Gold price, since it is assumed that investors in oil importing countries hedge inflation with gold (Investing.com, 2016)

·    USD / Rupee exchange rate (Investing.com, 2016),

·    WTI 1 month future, since if information on the market is perfect, that the derivative price is an unbiased estimator of the future price (U.S. Energy Information Administration, 2016),

·    US treasury bill rate (Quandl.com, 2016),

·    10 WTI return lags,

·    Google search terms (Google Trends, 2016),

o   Middle East

o   Oil price shock

o   OPEC

o   WTI

o   Terrorism

o   Fracking

The variable to be explained was the volatility in the next month, calculated as the standard deviation of the next 4 weekly returns of the WTI time series.

For the neural network, we split the data into 3 parts: Training (70%), Validation (15%), Test (15%).

Figure 2. shows the returns of WTI crude oil in the investigated period.

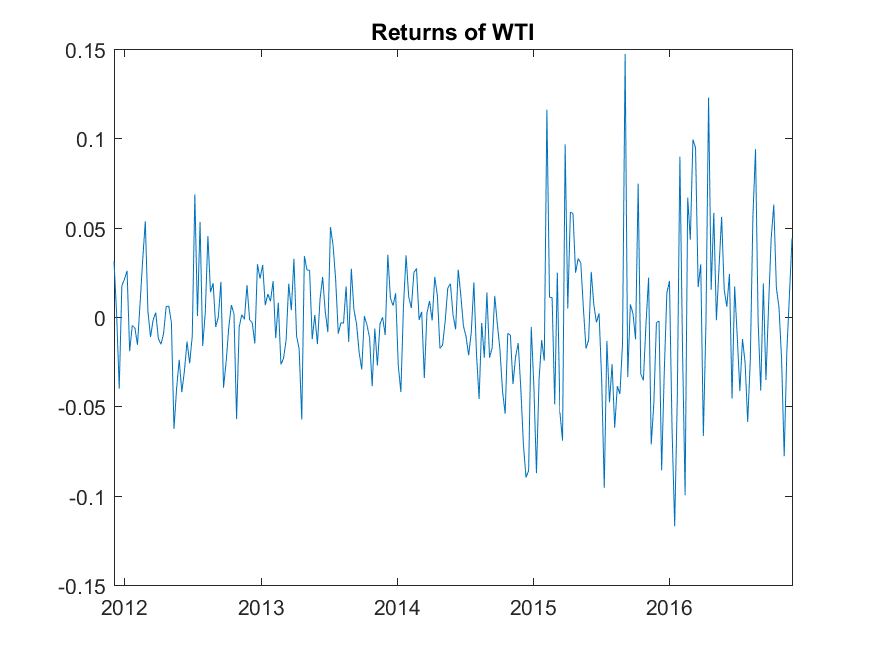


Figure 2 - Returns of WTI between 2012 and 2016  
Source: U.S. Energy Information Administration, 2016, n.pag.

Figure 3. shows the autocorrelation of oil prices in our sample.

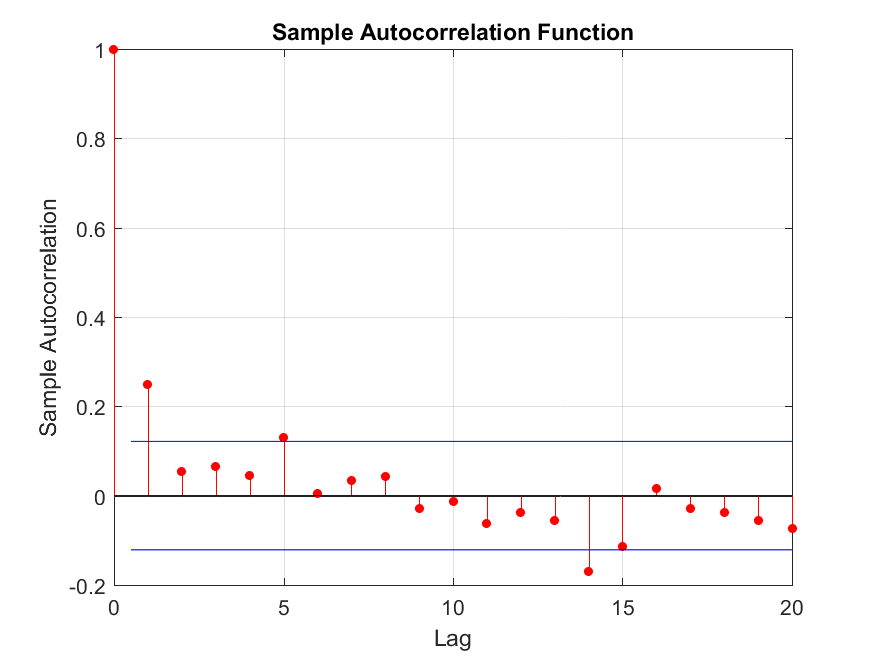


Figure 3 - Autocorrelation in the sample

Source: U.S. Energy Information Administration, 2016, n.pag.

## Regression methods

### ARCH

First of all, we tried to model the variance of the time series as an ARCH process. Basically, an ARCH model tried to model the variance with the previous error terms of the returns. For example, if the past 2 returns were unusually big, then an ARCH model would predict that the volatility is still going to be high in the next period, but not as big as previously.

We used a rolling window forecast. The ARCH model is estimated on the previous 10 observations, then we do a forecast for the variance of the next 4 weeks. These values are averaged, and we get the estimation for the average variance of the next 4 weeks, which is the target variable. This is done for every datapoint.

We estimated ARCH with various number of lags, but since  the return is only correlated with one previous lag, we decided to use the ARCH(1) model.

### Neural Network

As state in the previous part, we split the data into 3 parts, then trained the network using 5-fold crossvalidation. While searching for the optimal metaparameters of the model, we considered different number of neurons (1-9), different learning algorithms (Levenberg – Marquard, Bayesian), and also different learning parameters and the number of input variables.

We averaged most of these results to get the best estimation, and we only present the results of this final model, our best model. It is important to take the average not just the most accurate models, since it improves the generalization capability of the network.

### Neural network combined with ARCH

Finally, we also tried to improve the quality of the ARCH forecast by estimating the residuals of the ARCH model with a neural network. The idea here is that the ARCH model can capture the linear part of the volatility process, while the neural network will capture the rest.

# Results

## Fit to data

The following figures show the fit of each model to the original data. The blue line is the volatility of the data, the other lines are the models.

Figure 4 - Neural Network Fit to Data

Figure 5 - GARCH Fit to Data

Figure 6 - GARCH + NN Fit to Data

There are two conclusions to be drawn from the time series fit. One, the neural network could learn the datapoints where there was a huge shock, but could not grasp the small and intermediate volatility periods. Two, regressing the residuals of the ARCH produced a better fit than the ARCH itself. Unfortunately this will not improve predictive performance.

Figure 7 - GARCH and GARCH + NN Fit

Figure 8 - NN and GARCH + NN Fit

From the scatterplots we can see that our models underestimate the volatility in high volatility periods, but overestimate volatility in low-volatility periods.

## Predictive power

The results on the test set are summarized in Table 2. The bold RMSE scores correspond to the predictive performance of the neural network and the ARCH and the combined model as well. The non-bold scores correspond to the fit of the models to the data.

|  |  |  |
| --- | --- | --- |
|  | With Trends Data | Only financial Data |
| Train | 0,01312 | 0,01232 |
| Validation | 0,01051 | 0,00992 |
| Test | **0,01394** | **0,01500** |
| RMSEGARCH(0,1) | **0,01921** | **0,01921** |
| RMSEGARCH(0,2) | 0,01903 | 0,01903 |
| GARCH+NN Validation | 0,01710 | 0,01710 |
| GARCH+NN Test | **0,01882** | **0,01962** |

Table 2 - RMSE of different models on test set

The first observation regarding this table is that the neural network is a better predictor for the volatility for the next 4 weeks than any other model specified. Also, there is no significant difference between the two ARCH models. Third, the use of a neural network to predict the residuals of the ARCH models were a fruitless venture. It produced a better RMSE, but the improvement is not significant.

Fourth, and probably most important is that the fit of the neural network is 50% as good as its predictive performance. This indicates that we are dealing with an overfitted network.

Figure 9 - Predictions of NN, GARCH + NN compared to original data

Fifth, the role of Google Search data is not significant in our models. It improved the RMSE of the prediction, but the improvement is miniscule. This could be the result of an underlying overfit. We used too many variables for the network, and removing any variable from the dataset improved the forecast right away. On the other hand, it is possible that using only search terms as predictors might be better than using financial data for prediction. Therefore, our findings for the search hypothesis are inconclusive.

If we examine the above plot, another observation is evident as well: the neural network is good at estimating the volatility in high volatility periods, but not so good during regular times. This is because the additional information (Trends data) contains very little information about ordinary times, most of its predictive power comes from estimating big hypes.

Also, our models tend to underestimate the effects of big shocks, and overestimate the variance in quiet times.

## Further research

Our research did not have the necessary time and resources to investigate further possibilities for modelling.

* First, we did not employ any trend filtering methods for preparing our data. Possible options are using a simple moving average, or a Kalman-filter to deal with the trend and the noise in our data.
* Second, it is also possible to use a different machine learning model for the prediction, such as Support Vector Regression or decision tree. Using these methods would imply the application of ensemble methods as well. These methods could significantly improve the quality of our prediction.
* Third, it is possible to use the same methodology for different time series which are prone to volatility clustering and can be predicted by Google Searches.
* Fourth, we did not use any model comparison methods except for the RMSE. It would be possible to conduct a Diebold-Mariano test on the models, or to specify a different error function such as Mean Absolute Error. These methods could shed more light on the differences of our models
* Fifth, it is possible that the method works better if we change the frequency of our time series. Unfortunately, Google Trends data is only available in a weekly format for the last 5 years, and this fact imposes a serious limitation on the scope of our research.
* Sixth, we did not experiment with the selection of our predictor variables. When using neural network there is a big chance that the network will overfit, learning the data point and not the generalizable rule. Therefore it is not ultimately useful to use a lot of features, rather one should aim to select the best predictors for the model. We investigated the importance of Google Search, but not the role of the keywords themselves.
* Seventh, we need to investigate the optimal modelling complexity and procedure from the perspective of economic usability.

We plan to address (at least some of) the above topics in the Spring in a follow-up of our research.

# Conclusion

Our aim with the paper was to build on the existing research on the field of oil price and volatility forecasting. In order to achieve this goal, based on previous papers we have briefly summarized the factors behind oil price changes and volatility changes, showing that in contrast to the early years of forecast research, demand and supply are not the only determinants of oil price. Exogenous factors such as political conflicts and catastrophes, and endogenous factors such as speculators on the market also should be accounted for.  We have drawn up the landscape of previous research methodologies using econometrics and various machine learning methods, such as artificial neural networks or support vector machines.

Our hypotheses in the paper have been that artificial neural networks are better predictors of oil prices, than the standard ARCH model; that adding Google trends improves the previous forecast; and that running the neural network on the residuals of the ARCH model will provide a superior forecast to the previous models. Our results show that even though the model needs further improvements and ought to be applied on slightly modified datasets, the ANN did prove to be superior to the ARCH model, furthermore, adding Google trends data did slightly improve the RMSE of the model. However, applying ANN to the residuals of the ARCH did not yield any positive results.

Summarizing this paper, we see ample room for growth for this paper, as shown in our Further research section. Filtering data is a logical next step, as using weekly data only filters some of the noise. Different machine learning approaches, such as decision trees or SVMs can be utilized to compare results.

# Table of Figures

Figure 1 - WTI spot prices between 1986 and 2006, in USD per barrel Source: U.S. Energy Information Administration, 2016, n.pag. 5

Figure 2 - Returns of WTI between 2012 and 2016 Source: Andris ezt hivatkozzuk le! 13

Figure 3 - Autocorrelation in the sample 14

Figure 4 - Neural Network Fit to Data 15

Figure 5 - GARCH Fit to Data 16

Figure 6 - GARCH + NN Fit to Data 16

Figure 7 - GARCH and GARCH + NN Fit 17

Figure 8 - NN and GARCH + NN Fit 17

Figure 9 - Predictions of NN, GARCH + NN compared to original data 18

Table 1 - Major oil price shocks between 1973 and 2008, and their effect on production and prices. 7

Table 2 - RMSE of different models on test set 18

# References

Bénassy-Quéré, A., Mignon, V. and Penot, A. (2007). China and the relationship between the oil price and the dollar. In: Energy Policy, Volume 35, Issue 11, 2007, pp. 5795-5805.

Coppola, A. (2008). Forecasting oil price movements: Exploiting the information in the futures market. In: Journal of Futures Markets, Volume 28, Issue 1, 2008,, pp. 34-56.

Abraham, A. - Gabralla, L.A. (2013). Computational Modeling of Crude Oil Price Forecasting: A Review of Two Decades of Research. In: International Journal of Computer Information Systems and Industrial Management Applications, Volume 5, 2013, pp. 729-740.

Google Trends. (2016). *Google Trends. > Search keyword.* [Online] Available at <https://www.google.hu/trends/>, accessed on Dec 05, 2016.

Hamilton, J.D. (2009). Causes and Consequences of the Oil Shock of 2007-08. Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution, vol. 40(1 (Spring), pages 215-283. Available at <https://www.brookings.edu/wp-content/uploads/2016/07/2009a_bpea_hamilton-1.pdf>, accessed on Dec 11, 2016.

Hamilton, J.D. (2011). Historic Oil Shocks - Working paper. Prepared for the Handbook of Major Events on Economic History. Available at <http://econweb.ucsd.edu/~jhamilto/oil_history.pdf>, accessed on Dec 11, 2016.

Investing.com (2016). *Gold Futures – Feb 17.* [Online] Available at <http://www.investing.com/commodities/gold-historical-data>, accessed on Dec 05, 2016.

Investing.com (2016). *USD/INR – US Dollar Indian Rupee.* [Online] Available at <http://www.investing.com/currencies/usd-inr-historical-data>, accessed on Dec 05, 2016.

Kang, S.H., Kang, S.M. and Yoon, S.M. (2009). Forecasting volatility of crude oil markets. Energy Economics, [Volume 31, Issue 1](http://www.sciencedirect.com/science/journal/01409883/31/1), January 2009, pp. 119–125. [Online] Available at <http://www.sciencedirect.com/science/article/pii/S0140988308001539>, accessed on Dec 10, 2016.

Lizardo, R.A. - Mollick, A.V. (2010). Oil price fluctuations and U.S. dollar exchange rates. In: Energy Economics, Volume 32, Issue 2, 2010, pp. 399-408.

Morana, C. (2012). Oil price dynamics, macro-finance, interactions and the role of financial speculation. In: Journal of Banking & Finance, Volume 37, Issue 1, January 2013, pp. 206-226.

Moshiri, S. and Foroutan, F. (2006). Forecasting Non-Linear Crude Oil Futures Prices. The Energy Journal, Volume 27, Issue 4, 2006, pp. 81-96.

Pan, H., Haidar, I., Kulkarni, S. (2009). Daily Prediction of Short-Term Trends of Crude Oil Prices Using Neural Networks Exploiting Multimarket Dynamics. Frontiers of Computer Science in China. Volume 3, Issue 2, 2009, pp.177-191.

Quandl.com. (2016). *Treasury Bill Rates.* [Online] Available at <https://www.quandl.com/data/USTREASURY/BILLRATES-Treasury-Bill-Rates>, accessed on Dec 05, 2016.

U.S. Energy Information Administration. (2016). Weekly Cushing, OK WTI Spot Prices FOB. [Online] Available at <https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=RWTC&f=W> -> Petroleum & Other Liquids -> Data -> Crude Oil WTI Cushing, OK -> View History. Accessed on Dec 11, 2016.

Wang, F. - Wang, S. (2011). Analysis on impact factors of oil price fluctuation in China. In: Artificial Intelligence, Management Science and Electronic Commerce, 2nd International Conference. pp. 6146-6150.

Wei, Y., Wang, Y. and Huang, D. (2010). Forecasting Crude Oil Market Voltility: Further Evidence Using GARCH-class models. In: Energy Economics, Volume 32, Issue 6, November 2010, pp. 1477-1484.

Xie, W. et al. (2006). A New Method for Crude Oil Price Forecasting Based on Support Vector Machines. In: Computational Science - ICCS 2016. Springer Berlin Heidelberg.

Yu, L., Wang, S. & Lai, K.K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. In: Energy Economics, Volume 30, Issue 5, 2008, pp. 2623-2635.

1. Hamilton considers the Iranian revolution and the Iran-Iraq war as the causes of two separate price shocks close to each other in time. Production shortfalls mentioned in the table do not mean global supply decrease of the amount, as other oil producing countries have partly made up for the decrease. [↑](#footnote-ref-1)