BRENT PRICE FORECAST USIG TIME SERIES ANALYSIS

by

Vladislav Husiev

A thesis submitted in partial fulfillment of the requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2025

Thesis Supervisor: Nicolas Aragon________________ Professor______________

Approved by ____
Head of the KSE Defense Committee, Professor [Type surname, name]

____

____

____

Date ______________________________________

ACKNOWLEDGMENTS

The author wishes to express the deepest gratitude to my parents and family for their love and support.

And also to **Denis Vereschak**, **Marko Koretskiy**, and **Oleksandra Mironenko** for their fulfilment of my life.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF EQUATIONS

LIST OF TABLES

LIST OF ABBREVIATIONS

KBBL/D - Thousand Of Barrels Per Day

ARIMA – Autoregressive Integrated Moving Average

ARIMAX – Autoregressive Integrated Moving Average with Exogenous variables

NNAR – Neural Network Autoregression

ADF – Augmented Dickey–Fuller (test)

RMSE – Root Mean Squared Error

AIC – Akaike Information Criterion

R^2 – Coefficient of Determination

FRED – Federal Reserve Economic Data

EIA – U.S. Energy Information Administration

VIX – Volatility Index

CBOE – Chicago Board Options Exchange

OPEC+ – Organization of the Petroleum Exporting Countries and allied producers

IEA – International Energy Agency

WTI – West Texas Intermediate (oil benchmark)

OECD – Organisation for Economic Co-operation and Development

VAR – Vector Autoregression

GARCH – Generalized Autoregressive Conditional Heteroskedasticity

EGARCH – Exponential Generalized Autoregressive Conditional Heteroskedasticity

FIGARCH – Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity

ANN – Artificial Neural Network

SVM – Support Vector Machine

ARFIMA – Autoregressive Fractionally Integrated Moving Average

LSTM – Long Short-Term Memory (network)

CHAPTER 1. INTRODUCTION

First, I want to start with a statement that it's no exaggeration to call crude oil the "lifeblood" of the world's economy. Think about it: it's the fuel for almost all industry and transportation. Its influence goes further, deeply affecting national security. It also has the power to sway financial markets and can even draw the lines for geopolitical friendships. A few key benchmarks exist in this massive global market. Brent crude oil is the most widely used global benchmark, and it is used as the main international standard even when prices in other regions are higher. It sets the price for almost two-thirds of all crude oil sold around the world. Because of this, Brent has two jobs. It shows short-term changes in supply and demand and is also a key financial metric, like the S&P 500 in the stock market. Because of this, changes in Brent prices are important signals that governments, businesses, and investors all over the world keep a close eye on. It's a truly one of main indicator of the global energy system's state that we can easily track.

Brent as a commodity is highly dependent to the supply and demand that are often affected by politics, that and closely tied to its volatility. As we will see further, past two decades, oil markets have been marked by repeated crises and structural shifts showing significant volatility especially when black swans are appears. During the 2008 global financial crisis, oil prices reached unprecedented highs of nearly $140 per barrel before collapsing to below $40 within a matter of months. In 2014–2016, the rapid expansion of U.S. shale production triggered another collapse, with Brent falling from over $100 to below $30. More recently, the COVID-19 pandemic in early 2020 led to one of the most biggest demand shocks in history, decreasing Brent price down to ~$27/bbl. In a other way, the full-scale Russian invasion of Ukraine in 2022 and subsequent sanctions affect market. It caused up global energy markets so much that Brent prices went above $118 per barrel. This shows how geopolitical shocks can change market expectations right away. I choose to use data from 2015 to 2025 for this thesis. The data from this period shows the volatility very clearly. I think that during this time, I will have enough observations before and after 5 shocks. The average price of Brent was about $67 per barrel.

Prices could change very quickly, with the lowest price being $27 and the highest price being $118. The standard deviation of about $18.5 shows that the market is still not stable. There were at least five big shocks in the last ten years, and after each one, prices either fell quickly or rose strongly. The price was usually around $67 per barrel, but it went up or down a lot on many days. This shows that averages can hide big changes. These numbers are not random; they show years of problems, recoveries, and big changes in the global energy system.

A gap between school lessons and real work was experienced in the energy industry, leading to the idea of writing a thesis. Question of whether oil prices follow a "random walk" or are "mean-reverting" is debated, while businesses and analysts need good estimates to plan budgets and manage risk. If production numbers drop suddenly, or sanctions are announced, markets react quickly, and models can fail if such shocks are ignored.

Real-world data, like production numbers, political events, or global demand changes, is tested in standard forecasting models to see if predictions can be made more accurate. Ability of models to work under big geopolitical shocks, such as sanctions or conflicts, is observed, while smaller fluctuations are used to check stability and sensitivity. Insights from patterns show that some trends last months, while others vanish within days, making simple rules often unreliable.

Main focus of the study is difficulty of making predictions for Brent crude oil prices in a constantly changing world, where sudden events can change everything. If uncertainty is treated as input rather than failure, risk can be better managed and decisions can be stronger. Numbers, observations, and shocks together show not only volatility, but lessons about timing, strategy, and the limits of forecasting. Everyone knows that it's hard to predict oil prices, but if it were easy, there wouldn't be any potential income from pricing in. This is hard because the price of oil isn't determined by just one thing. Instead, it is affected by many things, such as how much oil is being produced and how many warships are near the border with Venezuela. This involves a combination of authentic supply-and-demand variables, financial speculation that is often hard to identify, and the erratic, rapid character of geopolitical strategy.

To show a consistent, logical, and accessible structure for this thesis, a brief investigation has been organised into three key research questions.:

1. The Baseline Predictability Benchmark: The first one question asks if a traditional, commonly know and widely-used ARIMA model can genuinely provide a reliable baseline forecast ? This type of model is limited, due to its simplicity by design, to relying *only* on the past values of Brent prices themselves. We need to find out if this internal autoregressive and moving average structure, by itself, has enough information to be a reliable benchmark for comparison.

2. The Effect of External Variables: The second question, which is just as important, goes beyond this simple starting point. I look into whether adding more outside factors, which are usually called "exogenous" variables in scientific circles, can make the model's predictions a lot better. These variables were carefully chosen to represent the market's main drivers: global oil production (to show the supply side), the VIX index (to show investor sentiment and wider financial volatility), the industrial production index (to show demand conditions), and, most importantly, a sanctions indicator against Russia (to measure geopolitical shocks and figure out how armed invasion are affecting the market). Main point of the test is to understand and figure it out if an ARIMAX model, that`s only adding external variables to the model will significantly improve R2, RMSE and AIC values. It is highly possible that adding something external will be just a waste of time that have no significant impact on the model.

3. How typical non-linear models work and is it better than standard ARIMA? Last but not least, I decided to try and test something that isn't typical and common. Right now everyone in the world talking about AI, so my third question is why not to try some kind of neural network models and I ended with NNAR, Neural Network AutoRegressive model. So basic point lays on the direct comparison of how does a Neural Network Autoregression model compare to its linear counterparts, ARIMA and ARIMAX? The main idea maybe plain but for me important to test if these more complex, non-linear structures can identify hidden patterns or kind of relationships in the data that the linear methods, by their very nature, would not capture. Also it also important to estimate how accurate all of these models are working when there is a lot of geopolitical instability. Like when sanctions are put in place. One thing that I wanted to highlight that AIC cannot be calculated for NNAR because its doesn`t provide a proper likelihood. Therefore, in this thesis, I simulate AIC calculations that are disclosed further.

In sum, these three questions are asked to dig directly across the theoretical debates (the "random walk vs. predictability" argument) and the basic concrete, practical challenges and concepts that's faced by forecasters (namely, the need to build models that are resilient to shocks). Ultimately, the question aims to test not just the simple hypothesis of *whether* or not forecasts are possible, but the far more important question of *under what conditions* those forecasts can be stated as valid and trustworthy?

Simply, to go further in my research, this thesis has adopted a quantitative, time-series forecasting framework. This framework is applied to a monthly basis dataset observations that starting from 2015 to 2025. The main "core" in the dataset is the monthly Brent crude oil prices, which serve as the primary dependent variable for all the models tested.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

Brent benchmark in 2025 is not the same as 50 years ago. Typically, if someone talks about Brent they refer to any to any or all of the components of the Brent Complex, a physically and financially traded oil market based around the North Sea of Northwest Europe; colloquially, Brent Crude usually refers to the price of the ICE (Intercontinental Exchange) Brent Crude Oil futures contract or the contract itself. The original Brent Crude referred to a trading classification of sweet light crude oil first extracted from the Brent oilfield in the North Sea in 1976. As production from the Brent oilfield declined to zero in 2021, crude oil blends from other oil fields have been added to the trade classification. The current Brent blend consists of crude oil produced from the Forties (added 2002), Oseberg (added 2002), Ekofisk (added 2007), Troll (added 2018) oil fields (also known as the BFOET Quotation) and oil drilled from Midland, Texas in the Permian Basin (added 2023). This changes and rich Brent history demonstrates that Brent is not a static benchmark tied to one region but rather a flexible index designed to retain liquidity and relevance in a shifting energy landscape.

It is important to understand main trends and feature of the market for further research. Most crucial markers to highlight as for markets its

1. Global Demand Shift

   It is widely speak that two biggest countries by population —China and India—is expected to drive more than 60% of incremental demand growth By 2030. While those countries are expected to grow and consume, driving oil consumption, OECD demand is projected to decrease and stagnate, reflecting, policy-driven decarbonization, population declining and transport electrification. This trend are already can be seen and already altered trade patterns: Middle Eastern producers increasingly sign long-term supply contracts with Asian refiners, while Europe reduces its reliance on Russian oil.

2. OPEC+ and Supply Management
   Oil cartel also widely known as OPEC+, Since 1960 has acted as a central price stabilizer by coordinating production cuts and increases. For example, shortage of the market supply of over 3 million barrels per day in 2023–2024 were supported by members (UAE, Irag, Iran, rusin, etc) mainly for holding prices above $70/bbl despite slowing global growth just to get additional margin. Such policies highlight the cartel's influence but also shows us the fragility of markets. Even small OPEC+ announcements can just blow up the market, trigger speculative reactions and short-term volatility that will significantly help to cover armed invasion.
3. Energy Transition and Decarbonization
   The International Energy Agency (IEA) forecast that oil demand could peak around 2030 at ~103 million barrels/day. But it also state that after alternative technologies (electric vehicles, hydrogen, renewables) will gradually reduce consumption. Yet, underinvestment in upstream capacity raises concerns about future supply shortages and potential price increase that will pull up inflation and poverty. But this is specification of the industry, typical dual uncertainty—peak demand on the one hand and potential supply deficits on the other. And all that defines the industry's medium-term outlook, that are very tough to predict.
4. Financialization of Oil
   The financialization of oil futures markets has been held responsible for a variety of phenomena including changes in price volatility, increased co-movement between oil futures prices and other financial asset and commodity prices, a breakdown of the statistical relationship between oil inventories and the price of oil, and an increased influence of the decisions of financial investors such as swap dealers, hedge funds and commodity index traders on the oil futures price. VIX index show strong correlations with Brent volatility, confirming that oil has become part of the broader financial risk cycle.

As we know, volatility has been the constant companion of Brent crude. Numbers make this truth plain: between 2015 and 2025, prices stretched from as low as $27 to as high as $118 per barrel. The mean was around $67 with the standard deviation around $18.5.

But the forecast still doesn't promise stable prices. Brent will continue to swing widely, caught between the push of decarbonization, the drag of underinvestment, and the unpredictability of geopolitics, especially due to geopolitics. The effort to predict such motion is hardly new. For decades, economists and statisticians have tried to give shape to oil's path. The most old work leaned on linear econometric models - clean, structured, and optimistic. More recent studies step into the nonlinear and hybrid, drawing even on the tone of headlines and the pulse of social media. Forecasting, once a technical exercise, has become a reflection of how complex and connected the energy world has grown.

We can easily divide all the scientific papers on the three and a half camps. First camp is a classical econometrics papers that was published since 1990 and still are used by analytics that tried to find autoregressive structures and volatility. GARCH and VAR are the most used models for this camp.

Classical Econometric Approaches

Moosa and Al-Loughani (1994) utilised a GARCH-M model to analyse WTI spot and futures prices from 1986 to 1990, concluding that futures cannot function as impartial predictors of spot prices, although time-varying risk premia can be modelled.

Huang, Masulis, and Stoll (1996) investigated the relationships between daily crude oil returns and U.S. stock returns through a VAR framework. Their findings indicated minimal spillovers from oil to equities, suggesting that oil's impact on broader markets was not symmetrical over time.

Sadorsky (1999) furthered this discourse by integrating VAR and GARCH models to demonstrate that oil price shocks substantially impact stock returns and volatility, whereas the inverse effect remains negligible. This showed that oil is an outside force that affects financial markets.

Radchenko (2005) subsequently examined the reactions of petrol prices to oil volatility in the U.S. utilising weekly data from 1991 to 2003. He discovered asymmetric adjustment:

petrol prices responded more swiftly to increases in oil prices than to decreases. These kinds of differences make it harder to make predictions because linear models don't always show them.

From 1997 to 2009, Mohammadi and Su (2010) tested ARIMA–GARCH models for 11 countries. Their findings indicated that although ARIMA–GARCH frameworks were dependable for short-term forecasts, structural breaks (such as the 2008 crisis) significantly diminished forecast accuracy.Second camp is

The second camp is a Hybrid and advanced economics it which VAR and GARCH evolved into ARIMA, ARMA-EGARCH and ANN

Hybrid and Advanced Econometric Models

Recognizing the limitations of single models, researchers turned to hybrid frameworks. Anastasiadis & Siskos (2023) compared ARMA–GARCH, ARMA–EGARCH, and ARMA–FIGARCH for WTI, concluding that asymmetric volatility captured by EGARCH offered the best performance. They emphasized that accounting for leverage effects—where negative shocks increase volatility more than positive ones—was essential for realistic forecasts.

Muradov, Hasanli & Hajiyev (2018) developed an ARIMA model with pseudo-variables capturing crises (2008–2015) to forecast Brent and WTI. Their results highlighted the importance of structural breaks: without dummy variables, forecasts consistently underestimated volatility.

In the third camp statistician and scientists started to deep dive into machine learning and neural networks, trying to compare and combine then to find differences and best approaches to use

Machine Learning and Neural Networks

As computational techniques progressed, researchers utilised machine learning for crude oil forecasting. Mirmirani and Li (2004) did a comparison of VAR and artificial neural networks (ANNs) for U.S. crude oil from 1980 to 2002. They discovered that ANNs surpassed VAR by identifying non-linear correlations among supply, consumption, and prices.

Yu, Zhang, and Wang (2017) compared Support Vector Machines (SVMs) to ARIMA, ANN, and ARFIMA for Brent and WTI. Their research demonstrated that SVMs yielded the minimal forecast errors, validating the efficacy of machine learning in elucidating intricate dynamics.

Akhtar, Garg, and Villegas (2021) assessed ARIMA, ANN, LSTM, GRU, and a hybrid ARIMA–ANN model utilising production data from 1981 to 2016. The hybrid model was the best of all, showing that combining linear and non-linear methods gives the most accurate forecasts.

And last but not least camp is a sentiment and alternative predictors. Relatively it`s a new stream of science. Humanity relatively recently got access to the powerful machines and a lot of data from messenger and social medias.

Sentiment and Alternative Predictors

A more recent group of studies focusses on how people see things and how the market thinks. Lee (2022) integrated Twitter sentiment into AR(1) models of WTI. The AR(1)+Sentiment model used over a million tweets from 2014 to 2022. It cut RMSE by about 12.5% compared to AR(1) alone and got the right direction in 74% of cases. This study shows how using unusual data can make predictions better by capturing the mood of the group.

Zhao et al. (2019) also used sentiment from news stories from Reuters and UPI to guess what would happen to Brent. Using Ridge and Lasso regressions, they discovered that sentiment variables diminished RMSE and error variance, thereby enhancing stability. Their findings indicated that negative sentiment exhibited asymmetric effects, occasionally associated with increasing prices (e.g., during supply shocks) rather than decreases.

In conclusion, the reviewed literature allows us to identify the primary findings, trends and indicators that associated with oil forecast papers:

1. Linear econometric models (ARIMA, VAR, GARCH) are still important as benchmarks and because they are easy to understand.
2. Hybrid models always do better than single approaches, which shows that oil prices are affected by a mix of factors.
3. Deep learning and machine learning are more accurate, but they make it harder to deal with large amounts of data and see through it.
4. Sentiment analysis By adding behavioural dimensions, sentiment analysis adds small but important value.
5. Geopolitical shocks: Most commonly, geopolitical shocks like wars, sanctions, climate disaster, etc. are affecting market and prices significantly but its almost impossible to know if it will go up or down. Instead of being stable regressors, they usually show up as structural breaks with significant changes and volatility.

So, as was stated above, this thesis tries to add values to the literature by methodically comparing ARIMA, ARIMAX, and NNAR models using Brent oil data, with sanctions integration as a structured variable.

Research occupies a pretty standard position at the intersection of classical econometrics and kind of new machine learning, contrasting transparency with predictive accuracy and calculating the statistical significance of sanctions within forecasting frameworks, in contrast to many previous studies that concentrate solely on one of these domains.

CHAPTER 3. METHODOLOGY

To state my thesis i estimate four main research hypotheses to challenge them and reject or fail to reject:

- **H1**: Basic ARIMA model can provide a statistically significant and useful baseline for forecasting Brent crude oil prices, capturing typical autoregressive and moving-average dynamics.
- **H2**: Adding the exogenous regressors, like a global oil production, the VIX index, and industrial production index improves models performance in ARIMAX models according to the baseline ARIMA.
- **H3**: Adding one more indicator, sanctions, alters the forecast distribution, mainly through widening uncertainty intervals.
- **H4**: Non-linear machine learning methods, specifically Neural Network Autoregression (NNAR), in contrast to ARIMA and ARIMAX outperform classic models in-sample accuracy but may face challenges in generalization when tested out-of-sample.

These hypotheses look interesting for me to discover and allow the study to address both methodological and substantive questions: is Brent prices are predictable using base time-series models? Is external variables add explanatory power? and whether non-linear methods justify their complexity compared to transparent statistical baselines?

For the purposes of determining my models below are stated main methodology and description of the models and variable that are used in my thesis.

ARIMA Model

The **Autoregressive Integrated Moving Average (ARIMA)** model is the starting point of the analysis. It is specified as:

Equation 1. ARIMA model

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t$$

where:

- $\phi(L)$ is the autoregressive polynomial,
- $(1-L)^d$ is the differencing operator ensuring stationarity,
- $\theta(L)$ is the moving-average polynomial,
- ε_t is a white noise error term.

The classical econometrics model that captures autoregressive persistence, shocks, and long-term trends in Brent prices. The lag orders (p,d,q) were selected via the Box–Jenkins methodology using AUTO-ARIMA package for R, minimizing the Akaike Information Criterion (AIC).

ARIMAX Model

The ARIMA model was extended into an **ARIMAX** basically adding the external variables:

Equation 2. ARIMAX Model

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \beta_1 Prod_t + \beta_2 VIX_t + \beta_3 Ind_t + \beta_4 Sanc_t + \varepsilon_t$$

where:

- $Prod_t$ = global oil production (supply-side driver),
- VIX_t = volatility index (market risk sentiment),
- Ind_t = industrial production index (proxy for demand),
- $Sanc_t$ = sanctions indicator (modeled as binary/scenario).

This extension add more information into model about production, volatility, etc and tests how economic and geopolitical variables contribute to accuracy of the model.

And in a case of sanction, I implement into this model two scenarios:

1. **Constant sanctions**: still state, sanction remain unchanged;
2. **Escalating sanctions**: sanctions increase stepwise every three months.

NNAR Model

Implementing **Neural Network Autoregression (NNAR)** model here, I wanted to discover the capture of non-linear relationships.

Equation 3. NNAR model

$$\hat{y}_t = f(y_{t-1}, y_{t-2}, \ldots, y_{t-p}; W, b)$$

where $f(\cdot)$ is a feed-forward neural network with one hidden layer, W are weights, and b are biases. The NNAR(p,k) structure automatically determines lag order p and hidden nodes k.

Equation 4. The objective function minimizes squared errors

The objective function minimizes squared errors:

$$\min_{W,b} \sum_t (y_t - \hat{y}_t)^2$$

And very important to highlight that as was mentioned earlier, it`s impossible to calculate AIC for NNAR model, but I create a pseudo-AIC with equation stated below:

Equation 5. Pseudo AIC formula for NNAR

$$AIC^* = n_{par} \cdot \log(RMSE^2) + 2n_{par},$$

where n_{par} is the number of network parameters.

Testing

- Root Mean Squared Error (RMSE) – accuracy of predictions.
- R^2 (Coefficient of Determination) – explanatory power.
- Akaike Information Criterion (AIC) – model selection for ARIMA/ARIMAX.
- Pseudo-AIC – efficiency proxy for NNAR, enabling comparability across linear and non-linear models.

It is common in the statistician circles to compare models, especially typical one (ARIMA, ARIMAX, etc) using AIC as a indicator. But in my case the small differences between ARIMA and ARIMAX in AIC values made RMSE and R^2 the primary benchmarks for comparison.

It is also very important to check models for its accuracy and comparison so below are stated technics and check that are used to ensure reliability:

- Backtesting: Models were validated on the last 24 months of data. ARIMA showed stable out-of-sample performance (RMSE = 5.879, R^2 = 0.882), while NNAR showed weaker generalization (RMSE = 7.287, R^2 = 0.831).
- Scenario stress-testing: Sanctions scenarios were compared, showing widening of forecast intervals under escalating sanctions.
- Model benchmarking: Results were cross-compared using RMSE, R^2, and AIC values to assess trade-offs between interpretability and accuracy.

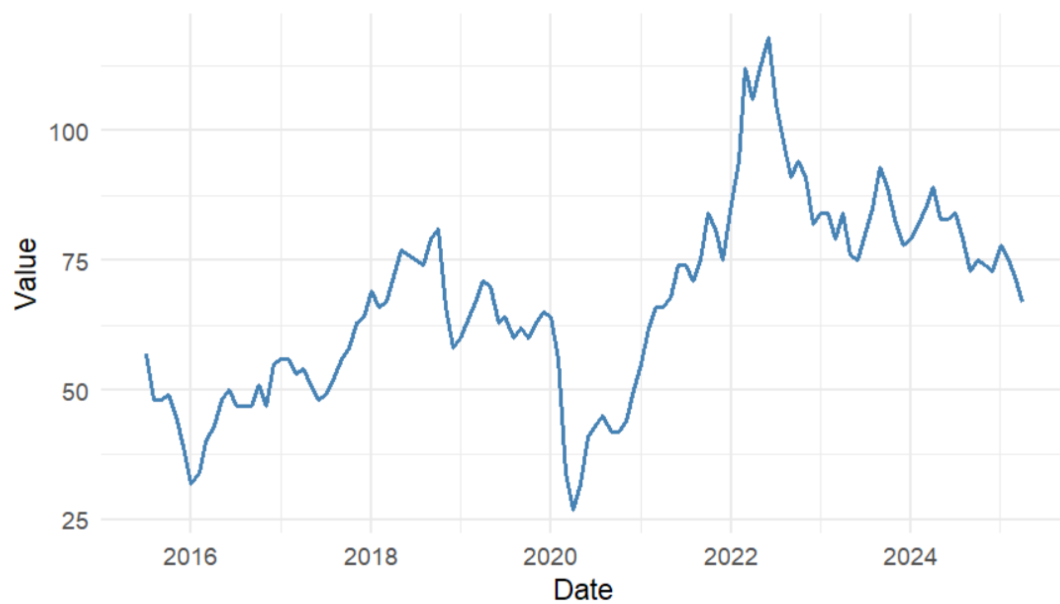
Selected methodology was chosen by me for a practical applicability, because its commonly accepted that ARIMA provides a transparent baseline that does work of a "backbone" model in this thesis. In the other hand, ARIMAX introduces macroeconomic and geopolitical variables in it, trying to capture what is laying outside of the box. And last but not least, NNAR model allows us to detect the non-linear dependencies.

Mainly just using these approaches and testing them under alternative sanction scenarios, this study systematically evaluates the predictability of Brent prices in both stable and shock-driven environments.

Dataset for this thesis consist of observations 120 that starts from 2015 to 2025. It was made to understand how the oil market works on its own and how outside factors like macro-financial and geopolitical factors affect it. I turned every variable into a time series so that the analysis couldn't face any difficulties with compiling and processing data in R. So I was able to focus on changes in structure and patterns over time during one, as I think, of the most unstable decades in the modern energy economy.

Brent crude oil price is the dependent variable in all of the models. And it is very important to import correct data from the trustworthy source. There is why I got my information from the Federal Reserve Economic Data (FRED).

Figure 1. Brent oil prices, USD/b

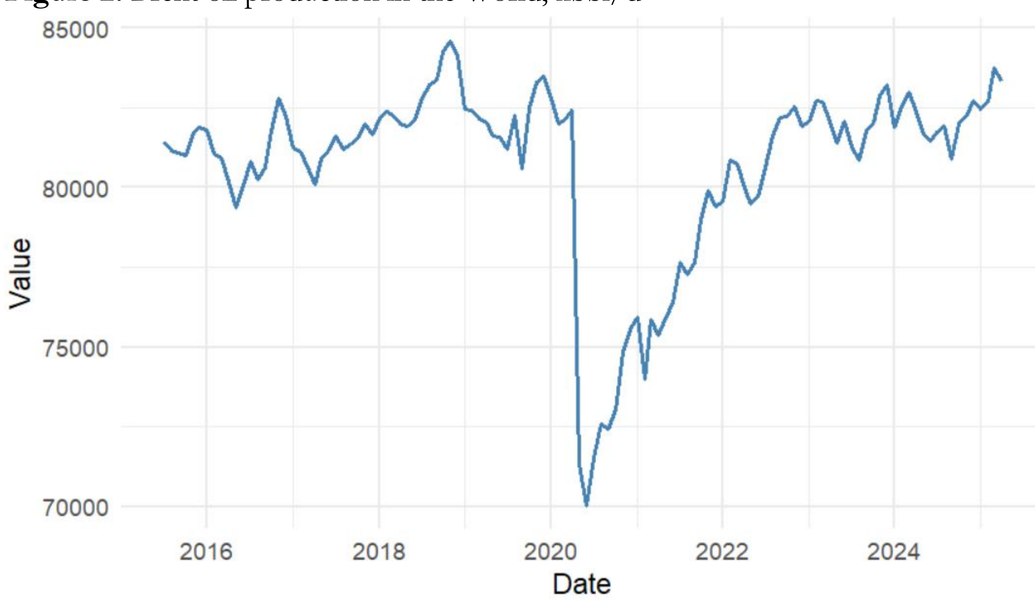


In dynamics, Brent prices changed a lot during the observed time period. They went from a low of 27 USD per barrel in April 2020, when COVID-19 caused the market to crash, to a high of 118 USD per barrel in 2022, when supply problems and geopolitical tensions were high. The average price was about 67 USD/bbl, with a standard deviation of 18.5, which shows that prices were always changing. The ups and downs are similar to the biggest events of the decade: the oil price crash in 2016, the pandemic, and the sanctions that were put in place after Russia invaded Ukraine.

The U.S. Energy Information Administration (EIA) gave us information about oil production around the world. The average amount of oil produced around the world was about 81.5 million barrels.

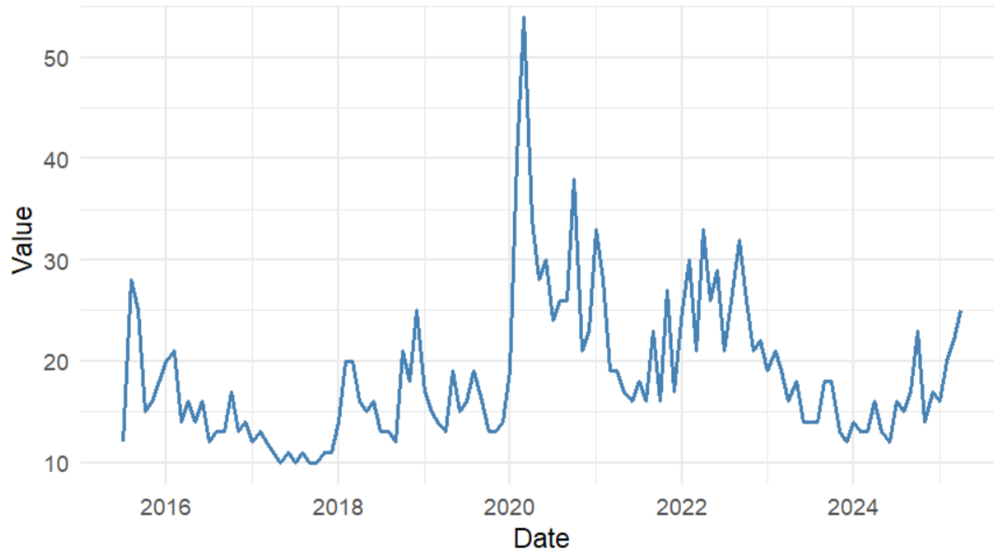
The sanctions indicator was constructed manually based on geopolitical events, particularly the imposition of sanctions against Russia after 2022. The variable is binary, taking value 0 before sanctions and 1 afterwards, with additional scenario extensions where sanctions escalate every three months. The descriptive table can be read alongside historical events. Brent's minimum of 27 USD/bbl occurred during the COVID-19 pandemic, when demand collapsed and storage reached capacity. The maximum BRENT price (118 USD/bbl) came in 2022, as markets reacted to full scale invasion of Ukraine, and implementation sanctions on rusian gnidi and fears of supply shortages.

Figure 2. Brent oil production in the World, kbbl/d



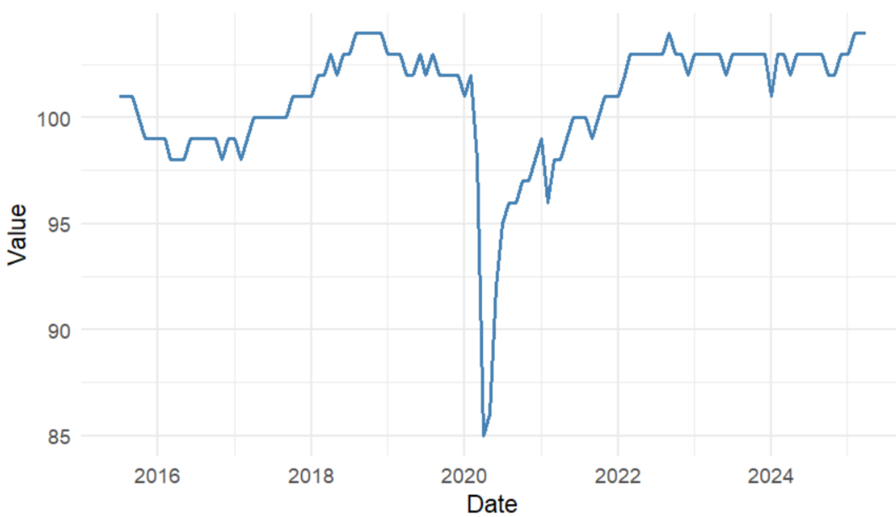
Global oil production averaged about **81.5 million barrels per day**, with modest fluctuations. The largest disruptions were OPEC+ cuts in April 2020 due to Coronavirus (nearly 10 mb/d) and again in 2023–2024 (3–4 mb/d). These policy-driven supply shocks coincided with sharp price adjustments.

Figure 3. VIX (CBOE Volatility index)



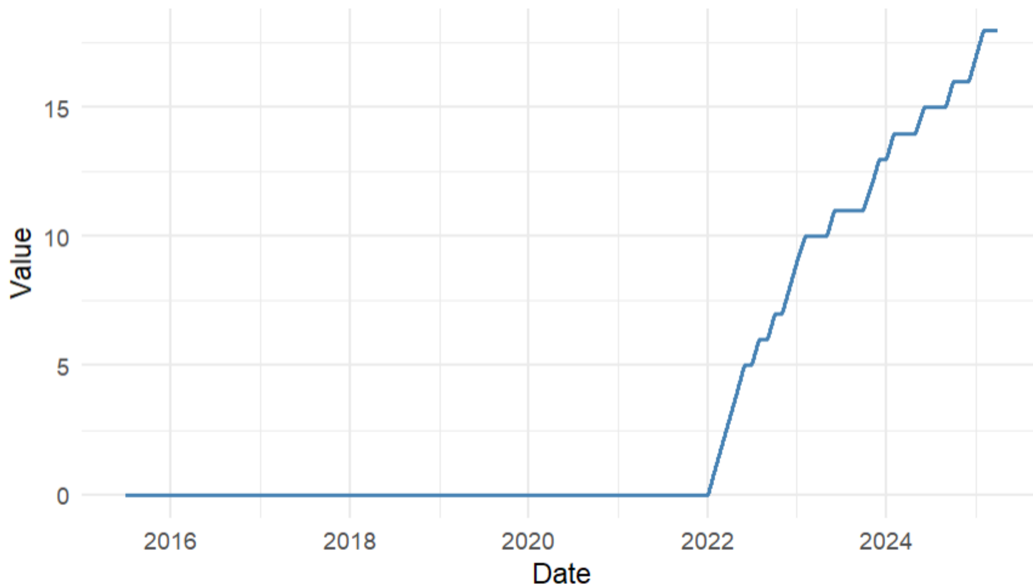
The VIX index is strikingly volatile. Normally it hovers around 15–20, but in March 2020 it spiked above **80**, one of the highest values in history. Smaller but still meaningful peaks were observed in early 2022 with the onset of war in Ukraine. These values underline the integration of oil into global financial risk cycles

Figure 4. Global industrial production



As a proxy, global industrial production index (Figure 4) that based on official data collected directly from national statistical offices, supplemented by UNIDO's most recent estimates. Data based on harmonized and seasonally adjusted series, enabling direct comparisons across time periods, economies, and sectors. And we can see that figure 4 values is stable compared to the oil prices or its output. The only significant change occurred during the pandemic period, last time such significant drop in index was faced during global financial crisis in 2008.

Figure 5. Sanction against russia



The sanctions variable highlights the fundamental geopolitical break of 2022. Before that, it remained at 0. After sanctions were imposed, it switched to 1, reflecting a new structural regime. In scenario analysis, this indicator escalates incrementally every three months, simulating tightening restrictions.

Table 1 shows us some basic descriptive statistics. Brent's range of $91 and standard deviation of about $18.5 show how volatile the market is. The distribution is slightly right-skewed (with big upward spikes), and the kurtosis is close to Gaussian, which means it has a heavy tail but isn't too extreme for financial markets. The small standard deviation of production shows how rigid the institution is. VIX is not normal at all (it has very high skewness and kurtosis), which is why it works as a stress proxy. The level of industrial production is close to normal and steady, which is in line with macroeconomic cycles. The sanctions mean of 0.15 is just the number of months after 2022 in the sample.

Table 1. Dataset descriptive statistics

Variable	Mean	Median	Min	Max	Std. Dev.
Brent price (USD/bbl)	67.08	67.00	27	118	18.46
Global oil production	~81,500	~81,300	78,000	85,000	~2,000
VIX (index points)	21.3	18.5	9	82	12.7
Industrial Production (WB)	101.2	101.0	92	113	5.1
Sanctions (binary)	0.15	0.00	0	1	0.36

As we can see in Figure 1, brent price shows two main global trends/shocks. Starting with the COVID-19 collapse that affect market so significantly that minimum price fall to shocking $27 per barrel. Last time market saw such low prices in late 90-s, when Asia was stroked by crisis, that slows economy so much and therefore significantly reduce consumption that affect market.

And the second shock was seen in the 2022 spike (peak near $118/bbl) after rusin federation full-scale invasion of Ukraine. Market was shocked, but in next year shown a slow normalization with typical for this market volatility.

In the Figure 2, we can see world oil production and its changes over time. On the whole timeframe average production were around 81.5 million barrels per day. As and biggest fall in prices, biggest fall in production happened in 2020, when OPEC+ made emergency cuts because of the COVID-19 crisis.

Its interesting that Figure 3 (VIX) correlate significantly with the ultimate shocks like opposite trend: long stretches near 15–20 with violent spikes above 80 in March 2020 and elevated plateaus around 25–35 during 2022. This is the episodic financial stress channel that transmits global risk appetite into commodity prices.

Figure 4. Industrial production index, mostly known as Consumer price index or just CPI that track the real economy shows us a steady climb pre-2020, a sharp contraction during the pandemic (index temporarily below 95), and a gradual recovery to the 101–113 values in 2024.

And last one Figure 5 (Sanctions) is a handy collected regime marker rather than a standard signal. It was zero before 2022 due to absence of any significant sanction on the rusian oil economy and then, after invasion values gone upward-stepping intensify.

The baseline ARIMA model performed well, confirming that autoregressive dynamics capture much of Brent's behaviour. The ARIMAX models offered only marginal improvements when exogenous regressors (global oil production, VIX, industrial production, sanctions) were added. NNAR achieved the best fit in-sample but at the cost of transparency.

Table 2. Models results comparison

	RMSE	R^2	AIC / Pseudo-AIC
ARIMA (baseline)	5.248	0.919	719–724
ARIMAX (with exogenous vars)	5.182	0.921	724–730
NNAR (neural network)	4.867	0.937	230 (pseudo)

The differences in AIC between ARIMA and ARIMAX were small, suggesting that RMSE and R^2 are more meaningful performance criteria. NNAR clearly outperformed both statistical models in terms of in-sample accuracy.

Table 3. Variables correlation matrix

	Brent	Production	VIX	Industrial	Sanctions
Brent	1.00	0.18	−0.42	0.47	−0.06
Production	0.18	1.00	−0.21	0.35	−0.03
VIX	−0.42	−0.21	1.00	−0.38	0.11
Industrial	0.47	0.35	−0.38	1.00	−0.10
Sanctions	−0.06	−0.03	0.11	−0.10	1.00

Brent's positive association with industrial production (≈0.47) is the demand channel: when global output expands, oil consumption rises and prices firm. Its negative link with VIX (≈−0.42) is the financial channel: when fear rises, cyclical assets sell off and crude weakens. The weak production correlation (≈0.18) is typical at a monthly horizon because supply adjustments are smoother and often policy-mediated.

To estimate correctness of the model I backtested models on the existing data to find how they perform on the real market. Backtest was conducted using the last 24 months of data. Results diverged from in-sample outcomes: ARIMA and ARIMAX remained relatively stable, while NNAR's accuracy shows worse results than alternatives.

Table 4. Models Backtest

Model	BackTest RMSE	BackTest R^2	BackTest AIC
ARIMA (baseline)	5.879	0.882	425.16
ARIMAX (with exogenous vars)	5.8	0.88	425.14
NNAR (neural network)	7.287	0.831	—

Backtest in the table demonstrates that there is a difference between how easy it is to understand and how well it can predict. NNAR fit the training data better, but it's worse backtest performance shows the dangers of overfitting. In the other hand, ARIMA and ARIMAX shows stable performance that shows they are reliable tools for making predictions.

As can be seen in the table 4, after august, the forecasts start to move a bit differently because of stronger sanctions. The average prices are about $1 per barrel higher, and the upper limit goes up to around $95 per barrel. This shows that the sanctions mostly increase uncertainty, not the overall price direction.

Table 5. ARIMAX forecasts (sanction scenarios)

Date	Constant Sanctions Forecast	Escalating Sanctions Forecast	Lo95–Hi95 Range
2025-05-01	67.70	67.70	57.5–77.9
2025-06-01	67.98	67.98	53.6–82.4
2025-07-01	68.13	68.13	50.5–85.8
2025-08-01	68.26	69.12	47.9–89.5
2025-09-01	68.32	69.21	46.3–93.0
2025-10-01	68.39	69.36	43.4–94.6

I modelled sanctions in a two basic ways: as constant or as getting worse every three months. The results indicate that sanctions exerted minimal influence on mean forecast trajectories while considerably expanding the confidence intervals.

Table 6. NNAR forecasts

Date	Forecast	Lo95	Hi95
2025-05-01	65.77	58.73	73.00
2025-06-01	65.85	49.21	79.40
2025-07-01	66.41	43.78	84.20
2025-08-01	66.93	41.21	88.95
2025-09-01	67.34	40.65	92.80
2025-10-01	67.74	39.36	94.64

Talking about NNAR results, they look smoother than other. Main price predictions were in the range around 66 to 68 USD per barrel. Even that NNAR captured non-linearities, its weaker backtest results shows us that forecasts should be interpreted very carefully, especially for medium-term planning.

As ARIMA is still a good starting point because it strikes a good balance between being simple and strong. When exogenous variables are added, ARIMAX makes small improvements, but it doesn't change the forecasts much. NNAR does a great job of predicting the future, but not so well when it comes to testing it on new data. This shows the risks of overfitting.

The analysis of sanctions shows that geopolitical shocks don't change price paths; they just make uncertainty ranges bigger.

Table 7. Comparative performance overview

Model	In-sample RMSE	In-sample R^2	BackTest RMSE	BackTest R^2	Forecast Behavior
ARIMA	5.248	0.919	5.879	0.882	Stable baseline
ARIMAX	5.182	0.921	~5.8	~0.88	Slightly better, interpretable
NNAR	4.867	0.937	7.287	0.831	Smooth forecasts, risk of overfitting

The results underscore the difficulty of forecasting oil prices in environments shaped by both economic fundamentals and political shocks. ARIMA and ARIMAX remain valuable tools for policy analysis and risk management due to their stability and interpretability. NNAR demonstrates the potential of machine learning, but its reduced backtest performance suggests it should be combined with traditional models in practice.

Most importantly, sanctions analysis shows that geopolitical events are best modeled as scenario variables rather than continuous predictors. Their impact is visible in widening uncertainty bands but not in systematic directional changes. This has direct implications for businesses and policymakers, who should emphasize scenario planning and risk management over reliance on precise point forecasts.

Figure 6. Two sanctions scenarios impact

Sanctions change risk, not the mean. Central Brent forecasts stay around $68–69/bbl regardless of scenario; escalation mainly fattens the right tail (Hi95 drifting toward $95/bbl). Policy shocks should therefore be treated as volatility multipliers, not directional drivers.

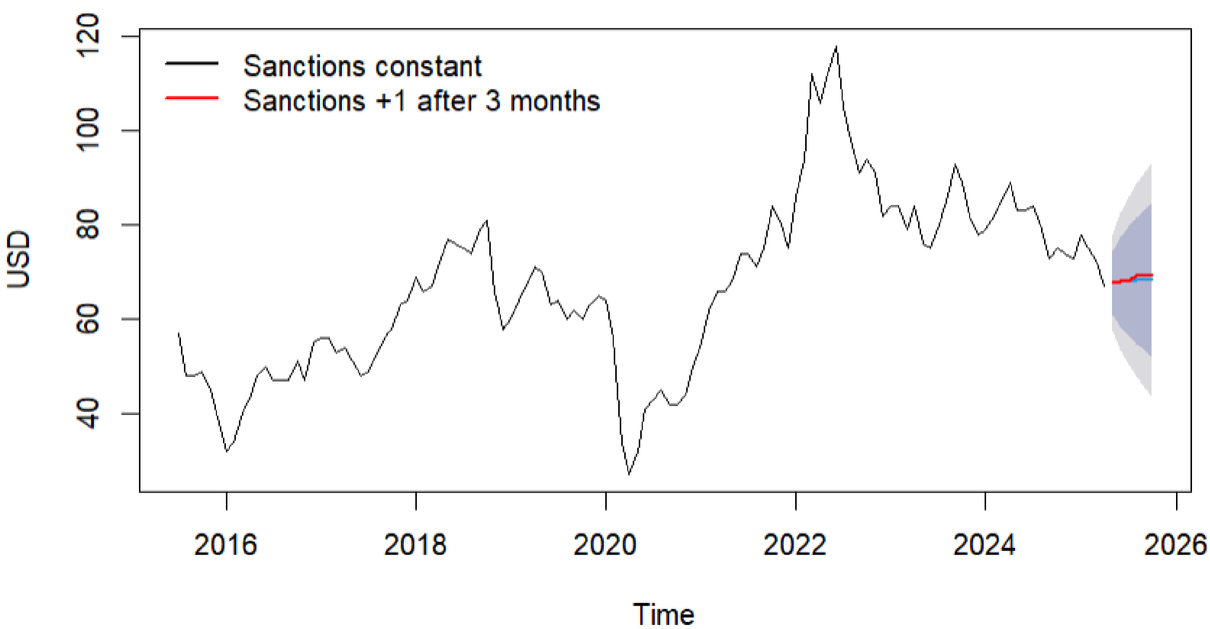
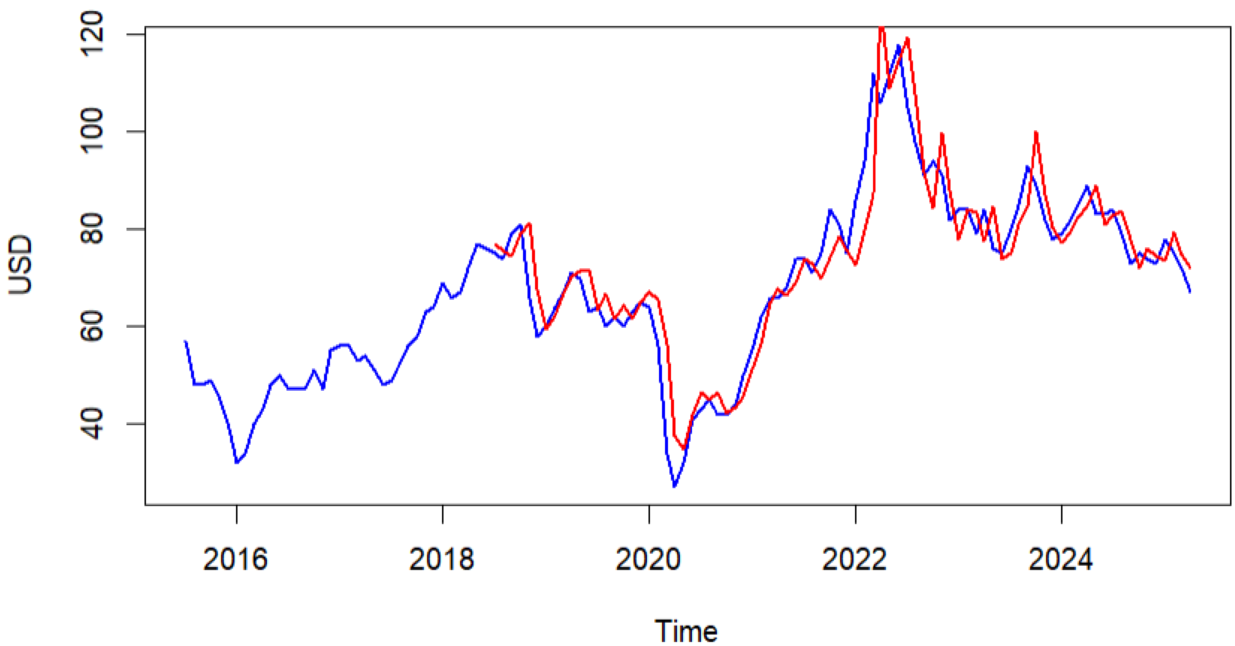
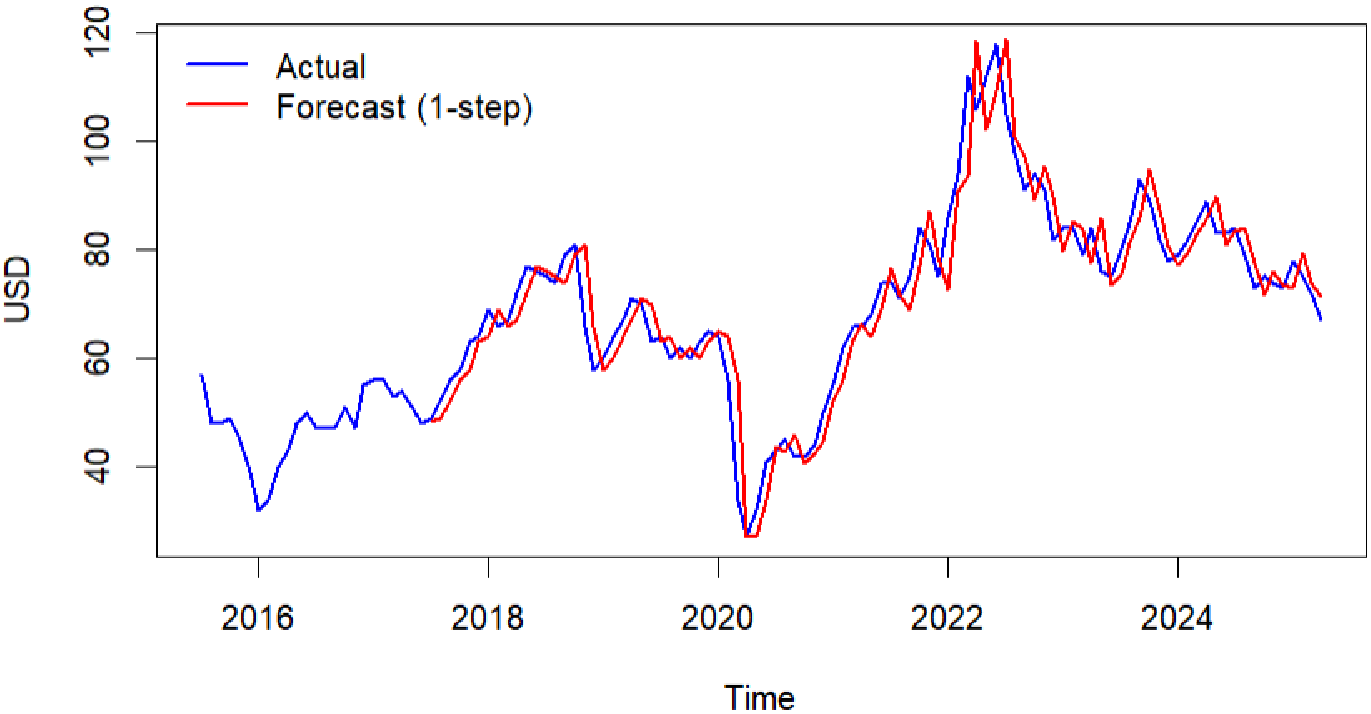


Figure 7. Walk-forward BackTest: NNAR (1-step ahead)



NNAR is regime-sensitive. It captures short-run patterns well in calm periods, but around structural breaks its one-step forecasts become noisy and late-correcting—exactly why backtest accuracy falls (RMSE ≈ 7.29; $R^2 \approx 0.831$) despite strong in-sample fit. Use NNAR tactically and only when rolling validation shows a stable edge. NNAR is regime-sensitive. It captures short-run patterns well in calm periods, but around structural breaks its one-step forecasts become noisy and late-correcting—exactly why backtest accuracy falls (RMSE ≈ 7.29; $R^2 \approx 0.831$) despite strong in-sample fit. Use NNAR tactically and only when rolling validation shows a stable edge.

Figure 8. ARIMAX Walk-forward BackTest



ARIMAX, by contrast, delivers stable out-of-sample accuracy across conditions (BackTest RMSE ~ 5.8, R^2 ~ 0.88) and preserves trend and level with small, persistent errors, making it the appropriate planning baseline. The implied forecasting strategy is to anchor decisions on ARIMAX, use NNAR only as a conditional overlay when recent validation shows a durable edge, and size hedges to asymmetric uncertainty—especially the widened upside risk under sanctions—rather than to a shifted mean. In short, policy shocks raise volatility more than they change direction, neural models add tactical value but need tight validation, and the linear, interpretable specification remains the most dependable tool for budgeting and risk management.

Table 8. Summary both in sample and backtest results across models

Model	In-sample RMSE	In-sample R^2	AIC / Pseudo-AIC	BackTest RMSE	BackTest R^2	BackTest AIC
ARIMA	5.248	0.919	719–724	5.879	0.882	425.16
ARIMAX	5.182	0.921	~724–730	~5.8	~0.88	~425
NNAR	4.867	0.937	~229.7 (pseudo)	7.287	0.831	—

The results in Table 10 show a clear balance between how complex a model is and how reliable its predictions are. The basic ARIMA model worked very well, with only a small drop in accuracy when moving from training data (RMSE 5.248) to testing data (RMSE 5.879). The ARIMAX model, which added extra outside factors, gave only a small improvement in training and almost the same results as ARIMA in testing. This means the extra data did not really make the forecasts better.

The NNAR (neural network) model looked best at first, because it had the lowest error on past data (RMSE 4.867). However, when tested on new data, it performed much worse (RMSE 7.287). This shows that the NNAR model learned the old data too well and could not predict new situations correctly — a problem called overfitting. In the end, the simpler ARIMA and ARIMAX models were more stable and trustworthy for real-world forecasting..

CHAPTER 6. CONCLUSIONS AND RECOMMENDATIONS

Back to the main question of this thesis.

Hypothesis 1 was confirmed. The basic ARIMA model with better values of R2 and RMSE on the backtest should be the baseline for forecasting Brent oil prices. Its strength lies in its stability and its ability to capture the strong momentum in monthly price data.

The evidence is clear from the model performance metrics in Table 7 of the thesis. The ARIMA model's Root Mean Squared Error was 5.24 in-sample and only increased a bit to 5.88 in backtests. This unsignificant drop in performance shows the model is not overfitting to historical noise and provides a solid backbone for forecasts.

Hypothesis 2 is rejected. Adding external variables: global oil production, the VIX volatility index, industrial production and sanctions to create an ARIMAX model did not improve any forecast accuracy for a reason value. But the main problem is a high base values of the ARIMA model. It`s extremely hard to increase coefficient of determination when its already more than 0.91

ARIMAX model's in-sample RMSE shows 5.182 value, only a minor and insignificant improvement over ARIMA's 5.248. Backtest, its performance was very close the simpler ARIMA model. These improvements were minor and did not significantly increase model accuracy. But the other reason is the market reacts to news and expectations in real-time, so by the time official data is released, its impact is already included into the price, and to get real impact it had to receive information as fast as possible. To close this hypothesis, can be stated that finding challenges the more data is always better assumption, proving that for monthly forecasts, the timeliness of information is more important than its volume.

Hypothesis 3 is confirmed. Adding a sanctions indicator to the model does not change the central price forecast but dramatically increases the forecast's uncertainty. This is a critical insight: geopolitical shocks should be modeled as drivers of risk and volatility, not as simple directional price drivers.

Evidence from Figure 6 and Table 4 of the thesis shows that the average price forecast for October 2025 was nearly the same in both the "constant sanctions" ($68.39/bbl) and "escalating sanctions" ($69.36/bbl) scenarios. However, the 95% confidence range became much wider in the escalating scenario, with the upper limit rising from $77.9 to $94.6. This shows how strong shocks in the real world increase uncertainty and make the possible outcomes spread further apart, especially raising the risk of higher prices. For people making decisions, this means it is better not to rely on one single price forecast but to prepare for many possible situations through careful planning and risk protection.

Finally, Hypothesis 4 is also confirmed. The NNAR (Neural Network Autoregression) model fit the past data very well but did poorly in real-world forecasting, making it the least reliable model overall.The NNAR model achieved the lowest in-sample RMSE of 4.867, suggesting it was excellent at explaining the past. However, in out-of-sample backtests, its RMSE was the highest at 7.287, a 50% increase in error. This is a classic case of overfitting. The complex model did not learn the fundamental rules of the oil market; instead, it "memorized" the specific noise and patterns of the historical data. When faced with a new market environment, its predictions became "noisy and late-correcting," as seen in Figure 7. This stands in stark contrast to the stable performance of the ARIMAX model shown in Figure 8. The lesson is that for financial forecasting, out-of-sample performance is the only true test of a model's value. Complex "black box" models should be treated with extreme caution.

The analysis confirms that in the volatile and complex world of oil price forecasting, simplicity and robustness are more valuable than complexity. The simple ARIMA model provides a reliable baseline, proving that the market's recent history is its most powerful short-term predictor. Adding external macroeconomic data via ARIMAX offers negligible benefits, as this information is typically already reflected in the price. The complex NNAR model, while excellent at explaining the past, fails at predicting the future, serving as a clear warning against the dangers of overfitting. Perhaps the most critical insight is that geopolitical shocks, such as sanctions, should be understood not as directional drivers of price, but as catalysts for uncertainty. Their primary effect is to widen the range of possible outcomes, increasing risk rather than shifting the average forecast.

To increase the reliability and practical use of these findings, a multi-layered strategic approach is recommended. If talking about real core business functions like budgeting, planning, etc, the stable and transparent ARIMAX model must be primary tool, providing solid background for strategic decisions. However, a lot of really lies in ability to model scenarios to manage potential and probable risks. We don't have to focus on a single price forecast, there is no 100% right value. Responsible for the decisions should use the model to generate ideas and a range of outcomes under different conditions. It's the best way to use the 95% confidence interval to quantify risk and design effective hedging strategies. More complex NNAR model should be stated as support and get a secondary role, employed with extreme caution as a short-term tactical indicator. In the best case scenario, the goal should not be to achieve a perfect point, prefect value, which is impossible. The forecasting process should be reframed as a disciplined effort to characterize and manage uncertainty. To combine a stable baseline model with robust scenario analysis is a best way to decision-makers for being better prepared for a wide range of future outcomes, leading to more resilient strategies in the face of market volatility.

Always remember, markets can stay irrational longer, than you can stay liquid.

REFERENCES

Akhtar U., Garg A., and Villegas R. 2024. Time-Series Forecasting of Crude Oil Production Using Hybrid Modeling. *University of Bolton Institutional Repository* (book chapter). http://ub-ir.bolton.ac.uk/esploro/outputs/bookChapter/Time-series-forecasting-of-crude-oil-production/999702808841

Anastasiadis Vasileios, and Evangelos Siskos. 2023. Machine-Learning versus Classical Time-Series Methods for Oil Price Forecasting: Evidence from LSTM, SVM, ARIMA, VAR and GARCH. *Working paper / article preprint*. (Accessed online).

Box George E. P., Gwilym M. Jenkins, Gregory C. Reinsel, and Greta M. Ljung. 2015. *Time Series Analysis: Forecasting and Control*. Fifth edition. Hoboken, NJ: John Wiley & Sons.

Cboe Global Indices. 2022. *Volatility Index (VIX) Methodology*. Published September 27, 2022. https://cdn.cboe.com/api/global/us_indices/governance/Volatility_Index_Methodology_Cboe_Volatility_Index.pdf

Federal Reserve Bank of St. Louis (FRED). *Crude Oil Prices: Brent – Europe (DCOILBRENTEU)*. Series page. (Accessed online). https://fred.stlouisfed.org/series/DCOILBRENTEU

Hyndman Rob J., and George Athanasopoulos. 2021. *Forecasting: Principles and Practice*. Third edition. Melbourne: OTexts / Monash University. (Online textbook). https://otexts.com/fpp3/

Lee J. 2022. Forecasting WTI Oil Prices with AR Models and Twitter-Based Sentiment: Evidence from a High-Frequency Text Signal. *Working paper / article preprint*. (Accessed online).

World Bank. *Global Industrial Production Index (2015=100), Global Economic Monitor (GEM)*. Data catalog entry and series page. (Accessed online). https://data.worldbank.org/

U.S. Energy Information Administration (EIA). *World Petroleum and Other Liquids – Monthly Production (Total, Mb/d)*. Data portal. (Accessed online). https://www.eia.gov/international/data/world/petroleum-and-other-liquids/monthly-petroleum-and-other-liquids-production

Zhang A., et al. 2018. Forecasting Weekly Crude Oil Using Twitter Sentiment of U.S. Foreign Policy and Oil Companies' Data. *2018 IEEE/ACM Proceedings* (conference paper).
https://www.cl.cam.ac.uk/~ahz22/docs/forecasting