NATIONAL BANK OF KAZAKHSTAN

Forecasting Oil Prices

Monetary Policy Department
Economic Study No.2024-6

Kulkayeva A., Taibekova A., Orlov K.

Economic research and analytical notes of the National Bank of the Republic of Kazakhstan (“the NBK”) are meant to disseminate the outcomes of the NBRK’s studies as well as of other academic research works of the NBK staff. Economic studies are disseminated in order to encourage discussions. The opinions expressed in the paper reflect the personal standing of the author and may not coincide with the official standpoint of the NBK.

Forecasting Oil Prices.

NBRK – WP – 2024 – 6

Predicting Oil Prices

Kulkayeva Altyn [1]
Taibekova Aida [2]
Orlov Konstantin [3]

Abstract

Several econometric models for forecasting oil prices are proposed in this paper. As a result, the developed models showed different forecast characteristics depending on the horizon. In the short-term forecasting period, good forecast properties were shown by autoregressive and moving average and vector autoregressive models with 5 lags, and in the medium term – by a vector autoregressive model with 13 lags. Combining the above models demonstrated the superiority of individual models in the short term (from 8 to 13 months). In general, it is recommended to use these models as an additional tool in designing the world oil price scenarios.

Key Words: oil, forecasting, combining, central bank
JEL-Classification: E32, E37, E59, Q43

[1] Kulkayeva Altyn – Chief Specialist-Analyst, Macroeconomic Research and Forecasting Division, Monetary Policy Department, National Bank of the Republic of Kazakhstan. E-mail: Altyn.Kulkaeva@nationalbank.kz
[2] Taibekova Aida – Chief Specialist-Analyst, Macroeconomic Research and Forecasting Division, Monetary Policy Department, National Bank of the Republic of Kazakhstan. E-mail: Aida.Taibekova@nationalbank.kz
[3] Orlov Konstantin – Deputy Head, Macroeconomic Research and Forecasting Division, Monetary Policy Department, National Bank of the Republic of Kazakhstan. E-mail: Konstantin.Orlov@nationalbank.kz

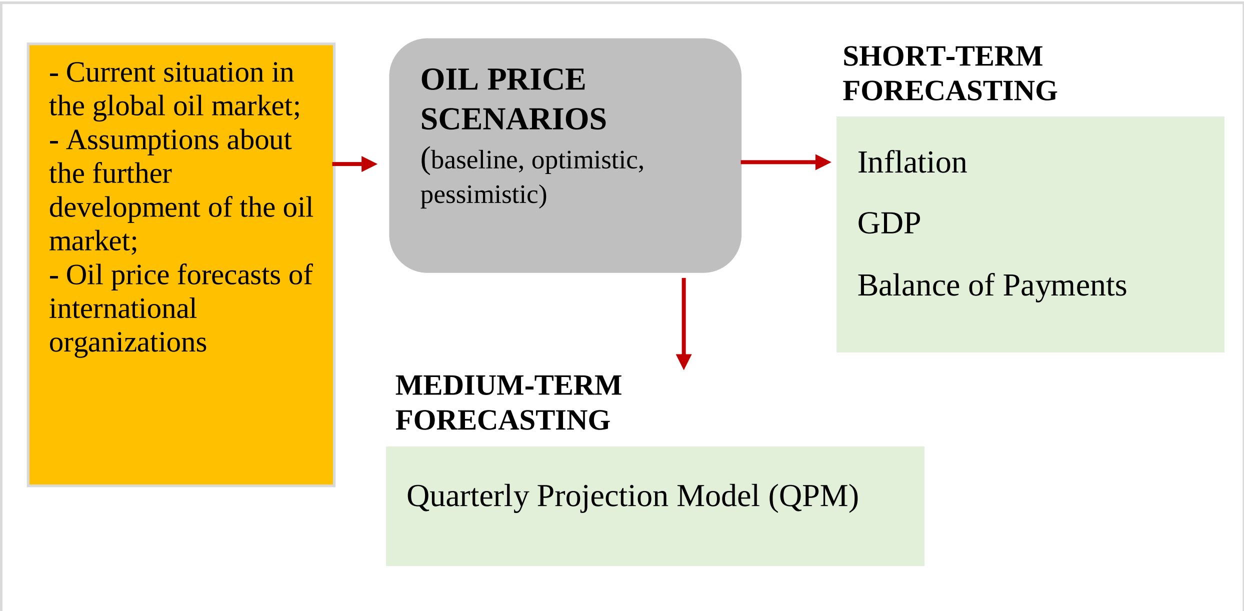
CONTENTS

1. PREAMBLE

In the modern world, oil remains one of the leading sources of energy. Fluctuations in its price can have both a positive and negative impact on individual countries or the global economy as a whole. From an economic point of view, a sharp increase in volatility in the oil market can lead to changes in macroeconomic indicators in both oil-exporting and oil-importing countries. In such conditions, forecasting oil prices is one of the effective solutions. Although forecasting oil prices still remains a difficult task since many factors have an impact on its dynamics, including geopolitical events, changes in oil supply and demand, natural and climatic changes, etc.

In the current practice of the National Bank, a scenario approach is used within the framework of the forecasting and policy analysis system (FPAS), which involves the development of the main (baseline) and two alternative scenarios for the price of oil (optimistic and pessimistic). In turn, the scenarios are designed by experts based on the analysis of the current and future situation in the global oil market, as well as taking into account the forecasts of international organizations (Consensus Ecs., EIA, IMF, OPEC, IEA and others). After approval by the Monetary Policy Committee, the oil scenarios are used in constructing short- and medium-term forecasts for inflation, GDP and other macroeconomic indicators (Diagram 1).

Diagram 1. Forecast and Policy Analysis System at the National Bank of the Republic of Kazakhstan



At the same time, the use of international organizations' oil price forecasts is associated with some difficulties. First, international organizations publish their forecasts at a frequency that does not always coincide with that required for preparing oil scenarios. Second, given high volatility of oil prices, current factors may not be taken into account in the forecasts published by international organizations. Third, the assumptions for these forecasts are often unknown, which may somewhat distort the overall macroeconomic picture.

In this study, Brent oil price forecasts were constructed for a horizon of up to two years (24 months). The main goal of the study is to compare the accuracy of various forecasting methods and identify the most optimal model for forecasting oil prices.

To achieve this goal, several econometric models and combined forecasts were constructed. The models were selected based on their wide application in econometric analysis and their ability to take into account various factors affecting the price of oil. A number of models were selected for the analysis: random walk models, autoregressive and moving average models, vector autoregressive models, and futures. The results showed that the autoregressive and moving average model has the best predictive properties only in the short term (up to 1 month). For the rest of the period, the vector autoregressive model demonstrates the best predictive qualities. At the same time, on a shorter horizon, the vector autoregressive model with 5 lags showed the best results, and the model with 13 lags – on a medium-term horizon. The use of conditional forecasting improved the results of the vector autoregressive model with 5 lags from 6 to 13 months.

The forecasts for these models were also compared with the forecasts of international organizations (quarterly forecasts). Thus, on the first quarter horizon, on average, a low forecast error was observed for Consensus Ecs., as well as for EIA and futures, and in the subsequent 7 quarters – for the vector autoregressive model with 5 lags and 13 lags. However, after excluding one dummy variable (March 2020), the forecast properties of the vector autoregressive model with 5 lags deteriorated sharply, and the model with 13 lags generally retained the quality of forecasts.

The combined model outperformed the results of individual models over a separate time span (from 8 to 13 months).

Based on the results of the study, it is recommended to use a set of models with the best result at each forecasting horizon as an analytical tool.

2. LITERATURE REVIEW

Oil price forecasting has become increasingly popular since the global financial crisis and it includes various methods. The main sources of forecasts for central banks are the International Monetary Fund (IMF), Consensus Economics, the US Energy Information Administration (EIA) and other international organizations, which generally use survey, econometric, financial, fundamental or combined forecasting mechanisms. Some central banks additionally use their own forecasts (Russia, Azerbaijan).

The IMF uses futures markets for its annual forecasts and publishes oil price forecasts in its World Economic Outlook reports. The Consensus Economics forecast is an average of the opinions of various commercial banks and investment houses. The EIA presents monthly forecasts of the spot price of Brent crude oil.

EIA uses three sources for its projections:

1. A combined model where the final forecast is generated as the mean of forecasts by five different models:

1) vector autoregressive model - VAR (oil production, the US CPI-deflated cost of US refineries purchasing imported crude oil, change in world oil stocks);

2) a model based on the spread between futures and spot prices;

3) a model using prices for other types of primary materials except oil;

4) a model with a changing over time parameter that reflects the ratio between (a) a spread between the US spot prices for gasoline and fuel oil and (b) a spot price for crude oil;

5) a model based on cumulative changes in the US crude oil reserves;

2. A linear regression where own forecasts of independent variables are used: a monthly change in the US oil inventories, total oil inventories in the OECD member countries relative to the prior four years' average for each month; a monthly change in global GDP;

3. Expert judgments based on the understanding of global oil market and the forecast of the balance of demand and supply.

Based on the results of the first two models, EIA analysts prepare final forecasts. The first methodology of five separate models is discussed in the study by Baumeister, Kilian, Lee (2014). In particular, the work confirms that combining the models yields better results in predicting oil prices than EIA's own forecasts.

In general, most of the research in oil price forecasting, oil price fluctuations, the role of speculation in the global oil markets, and other aspects of the oil market belongs to Lutz Kilian. Most of the researchers' work is based on his 2009 study on identifying structural shocks in the price of oil.

The vector autoregressive (VAR) models developed by Kilian (2009) have been used by other researchers for short- and medium-term forecasting of real oil prices. Until then, practitioners had long relied on futures. The topic of forecasting was further developed in his joint works such as Baumeister, Kilian (2011) and Alquist, Kilian, R.Vigfusson (2013). In particular, they showed that predictions of real oil prices based on VAR models provide more accurate forecasts compared to futures or other models.

Other studies on forecasting have experimented with different specifications of variables, a different set of variables, and a combination of models. The work of the ECB specialists – Manescu C., Robays I.V. (2014) – demonstrated that a combined model (4 models: futures, risk-adjusted futures, BVAR, DSGE) showed more accurate and robust forecasts of quarterly real oil prices than individual models. The work of IMF specialists Beckers, Beidas-Strom (2015) used such variables as global industrial production, oil supply, inventories, inflation, the dollar exchange rate, trade weighted by large consuming countries. The study used such data as world and regional production indices, oil demand in developed and emerging markets, supply in OPEC and non-OPEC countries. As a result, the authors come to the same conclusion that the standard VAR model shows better results compared to futures and random walk. Combining forecasts yielded better results over 18-month horizons.

In addition to classical econometric approaches to time series analysis, machine-learning methods (support vector machines, neural networks, long short-term memory networks, etc.) or their hybrid versions with basic models have been used in recent years to forecast oil prices. These studies[4] show superiority of such models compared to classical methods. The advantage of neural networks is that the model can continuously track the volatile dynamics of crude oil prices and model non-linear and complex relationships.

Thus, a review of the available literature makes it clear that even a more accurate model cannot consistently produce good forecasts. Some models may produce good results in the short term, while other – in the medium term. Overall, the authors note that predicting oil prices based on a long data sample with structural breaks remains a difficult task.

3. DATA

The following variables have been used in this study:

Category	Notation	Definition	Unit of Measurement	Source
Crude oil price	brent	Brent oil price	US dollars per barrel	EIA
Real oil price	brent_real	Brent oil price deflated by the US inflation	US dollars per barrel	EIA
Oil supply	oil_prod	global oil production	mln barrels per day	EIA
Oil stocks	stocks	world oil inventiries (in the OECD member countries)	mln barrels	EIA
Proxy for oil demand	rea[5]	Index of global real economic activity	gap from a trend	Federal Bank of Dallas
Macrofactors	cpi_accum	US inflation	consumer price index	US Bureau of Labor Statistics
Financial	futures	futures	US dollars per barrel	Refinitiv (Thomson Reuters)

Monthly data from January 2003 to December 2023 (252 months) were used because crude oil inventory data is only available from this period. For earlier periods, oil inventory data is only available in aggregate form. In general, EIA’s oil inventory data include the OECD inventories (about half of world stocks).

---

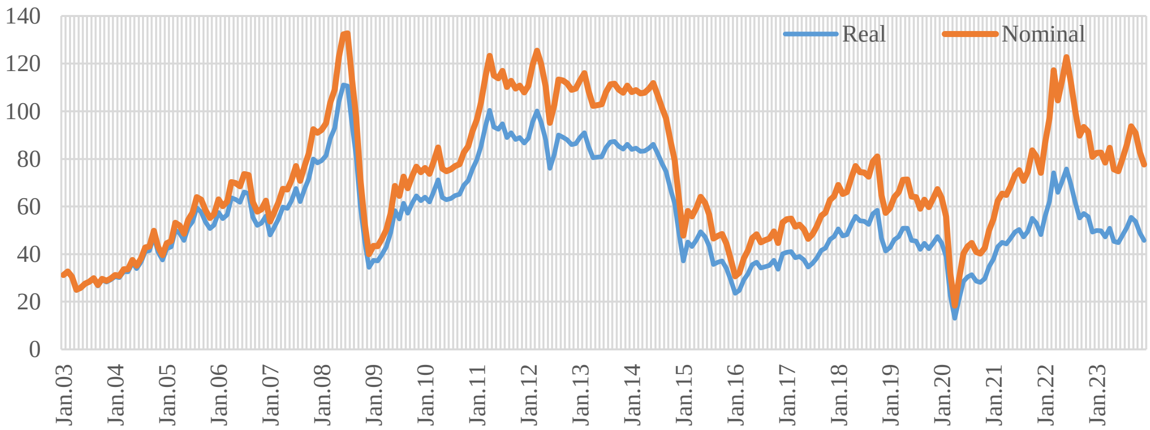
[4] Abdollahi, H., & Ebrahimi, S. B. (2020), Lee Jo Xian и др. (2020), N.Gupta, Sh.Nigam (2020), H.Alrweili and H.Fawzy (2022). K. Zhang (2022).

[5] Index of global real economic activity in the markets of industrial production as proposed by Kilian (2009) with an adjustment discussed in Kilian (2019). This business cycle index is expressed as a percentage deviation from trend. It is calculated on the basis of global dry cargo freight rates denominated in the US dollars and can be considered as an indicator of the volume of freight in the global industrial production markets. This index is updated monthly.

Brent crude oil data were chosen because Brent is a benchmark that is more commonly accepted as a global benchmark and is the basis for developing scenario-based forecasts of macroeconomic indicators. It is worth noting that Kilian (2020) recommended using the cost of purchasing imported crude oil by US refineries for the global oil market as a proxy for the global crude oil price. He did not use Brent because the historical series dates back to the mid-1980s (in his work, oil price data dates back to 1973). Some authors still forecast Brent for the above reasons.

Monthly data are used because the monthly path of oil price is important for its subsequent use in developing short-term forecasts of macroeconomic indicators in the National Bank. For forecasting purposes, the nominal price of Brent crude oil was converted into real terms using the US inflation (Graph 1).

Graph 1. Brent Oil Price



Sources: EIA, US Bureau of Labor Statistics, authors' computations

Note: Nominal Brent price deflated by the US CPI base index (base = December 2002) represents the real oil price.

All variables were brought to stationarity by taking logarithms, except for the Kilian index. The Kilian index variable is initially stationary by the type of construction, since it is a gap from the trend.

The variables of oil prices, stocks and oil production are reflected in growth rates (in the first difference of logarithms). According to the results of the unit root test, these variables showed stationarity at the first difference with the 5% confidence level. At the same time, the oil price at the level turned out to be stationary at the 10% confidence level (the Dickey-Fuller test showed non-stationarity of the oil price variable at the level with the 5% confidence level (p-value=0.0592) and stationarity in the first difference (p-value=0,000)).

A recurring question among applied researchers is whether real oil prices should be reflected in the oil market models in logarithmic or growth rates. The literature on forecasting real oil prices often suggests that logarithmic autoregressive models tend to be more accurate than the growth rate models.

However, in our case, the model with oil variables in logarithms and growth rates yields different forecasts, so the final specification of the variable was model-specific.

No seasonality was detected in the variables, so they were not seasonally adjusted.

All computations in this study were performed by using the Eviews econometric software package.

4. MODELS FOR FORECASTING

Random walk, ARMA

The autoregressive moving average (ARMA) model and the random walk model were chosen as the basic benchmark models due to their suitability for comparing forecast results with other models.

During the construction of the models, the use of real oil price in the first difference significantly worsened the forecast, while the use of logarithms had the opposite effect. Taking this into account, as well as the results of automatic selection of the autoregressive model, it was decided to use the natural logarithm.

The selection of the optimal ARMA model was carried out using the auto.arima() function. As a result of the automatic calculation of 169 models, the ARMA model with parameters (5,2) was selected as the most optimal. For comparison, an ARMA model with parameters (1,1) was also built. In the work of Benjamin Beckers (2015), the same model was chosen as a benchmark. Both models, when evaluated on the test part of the sample (from January 2021 to December 2023), showed good predictive properties. The root-mean-squared error for ARMA (1,1) was 12.38, for ARMA (5,2) – 13.12. In comparison with the actual values of oil price, ARMA (5,2) captured all points of price growth and decline almost across the entire forecast horizon, while ARMA (1,1) reflected only the general trend.

The random walk (RW) model assumes that the predicted price of oil in the future will be equal to its current price ($\hat{y}$ = y_{t-1}). In this paper, we used the random walk model without drift.

Futures

According to this approach, the forecast of oil price for the period *h* is the price of an oil futures contract with maturity at a certain date. Futures contracts are financial instruments that allow traders to fix the price today at which to buy or sell a certain amount of a commodity on a predetermined date in the future. Many central banks and international organizations use this approach, since it is a simple and convenient tool, which can provide information about market expectations. However, futures prices are not equal to the expected oil price; they can deviate from the spot price due to the risk premium component.

Vector Autoregressive Model (VAR)

In a generic form, the model is presented as follows:

$$y_t = \sum_{j=1}^{p} A_j \cdot y_{t-j} + B_t x_t + C d_t + \varepsilon_t, \qquad (1)$$

where y_t – k-dimensional vector of endogenous variables, x_t – vector of exogenous variables, d_t – vector of dummy variables, A_j, …, A_p, B_t, C – matrices of coefficients to be estimated, and ε_t – vector of residuals.

The approach to developing the model was based on the study by Baumeister, Kilian (2014). The original work uses a structural VAR of four variables to identify demand, supply, and speculative shocks. However, such model in its standard form can also be used for forecasting. Four variables were used to build the model: the real price of Brent crude oil, world production, oil stocks, and the real global economic activity index.

To reflect the crisis moments in oil prices, a dummy variable was introduced, taking the value of 1 in November 2008 (the global financial crisis) and March 2020 (the pandemic). Adding a dummy variable improved the forecasting qualities of the vector autoregressive model. When determining the optimal lag, the Akaike information criterion (AIC) tests showed a lag significant at the level of 13 in the model using rates and 5 in the model using logarithms. In this regard, two models were selected: VAR5 with the logarithm of oil prices and with the inclusion of a dummy variable, VAR13 in the rate of growth of oil prices (dlog) with the inclusion of a dummy variable.

In the study by Kilian, Baumeister, 12 lags were used. The authors themselves noted that models with short lags cannot capture slow cycles of decline and will underestimate the importance of demand shocks. In this regard, it was recommended to focus on longer lags.

Bayesian Vector Autoregression Models (BVAR)

To solve the problem of overparameterization or the “curse of dimensionality”, Bayesian vector autoregression (BVAR) was used in this study. In order to “compress” the model parameters, the Minnesota prior distribution proposed by Litterman (1986) and Doan et al. (1984) was used. Considering that all the variables used are stationary, the parameter μ was taken equal to zero. The remaining hyper parameters, based on the results of searching for optimal values, were taken in the following values: $\lambda 1 = 10$, $\lambda 2 = 0.3$, $\lambda 3 = 0.1$. The number of lags was kept similar to the vector autoregression – 13 lags. However, in comparison with the vector autoregressive model, the exogenous parameter of the dummy variable was excluded due to the deterioration of the predictive power of the model.

Based on the obtained results, the dynamics of the mean square errors over almost the entire forecast horizon turned out to be similar to the VAR results. In this regard, it was not used later.

5. FORECASTING RESULTS

Assessing Forecasting Properties of Individual Models

To better illustrate the change in forecast performance over time, we use a rolling window method instead of a recursive approach, which would less clearly reflect the instability of forecast accuracy.

Within each model, after estimating its parameters on a rolling sample, forecasts were made from 1 to 24 months[6], and forecast errors were calculated. The first forecast was obtained for the period January 2018 – December 2019, then the

[6] In the absence of actual data for comparison with the forecast, the horizon was obviously shortened.

sliding window for estimating the model in 180 months was shifted by 1 month forward, and a forecast was made again for the next 24 months. Thus, the forecast properties of the models were estimated based on 72 rolling samples, and for each forecast month, the root-mean squared error of the RMSE model was calculated for all rolling samples. Table 1 presents the root-mean squared errors for all estimated models for 24 months.

Table 1. Root-Mean Squared Error (RMSE) of the Forecast for Various Models Relative to RW

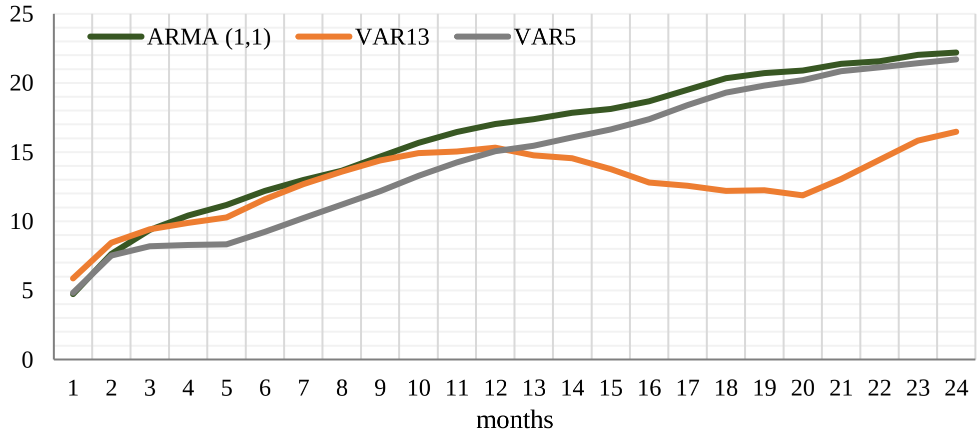
Model \ Month	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
ARMA (1,1)	**0.97**	0.99	1.00	0.99	0.98	0.98	0.98	0.96	0.96	0.95	0.95	0.95
ARMA(5,2)	2.00	1.27	1.17	1.06	0.98	1.17	1.10	1.09	1.01	0.95	1.00	0.97
VAR5	0.99	**0.98**	**0.88**	**0.81**	**0.75**	**0.76**	**0.79**	**0.81**	**0.81**	**0.83**	**0.84**	0.86
VAR13	1.20	1.10	1.00	0.94	0.90	0.93	0.96	0.96	0.94	0.91	0.87	**0.85**
Futures	1.77	2.18	2.31	2.39	2.36	2.30	2.29	2.24	2.18	2.14	2.12	2.13

Model \ Month	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
ARMA (1,1)	0.94	0.94	0.93	0.93	0.92	0.93	0.93	0.92	0.92	0.92	0.92	0.92
ARMA(5,2)	0.96	0.94	0.93	0.98	0.94	0.93	0.92	0.92	0.93	0.92	0.91	0.91
VAR5	0.86	0.88	0.89	0.89	0.91	0.92	0.92	0.92	0.93	0.93	0.91	0.91
VAR13	**0.80**	**0.77**	**0.71**	**0.63**	**0.59**	**0.56**	**0.55**	**0.52**	**0.56**	**0.62**	**0.66**	**0.68**
Futures	2.13	2.12	2.11	2.08	2.05	2.03	2.05	2.05	2.03	2.02	2.00	1.98

Note: In the table, a mean squared error value of less than one indicates that the model outperforms the Random walk results in the specified period. Values in bold indicate the superiority of the model compared to the other models presented.

According to the results obtained, on a one-month horizon, the relative forecast error is the smallest for ARMA(1,1). Starting from the second month and almost up to a year, VAR5 demonstrated good forecasting properties. At a later period, namely from the 12^{th} month until the end of the forecast horizon, the VAR13 model became the "champion" (Graph 2).

Graph 2. Dynamics of Mean Squared Errors in Models with Better Predicting Properties



Source: authors' computations

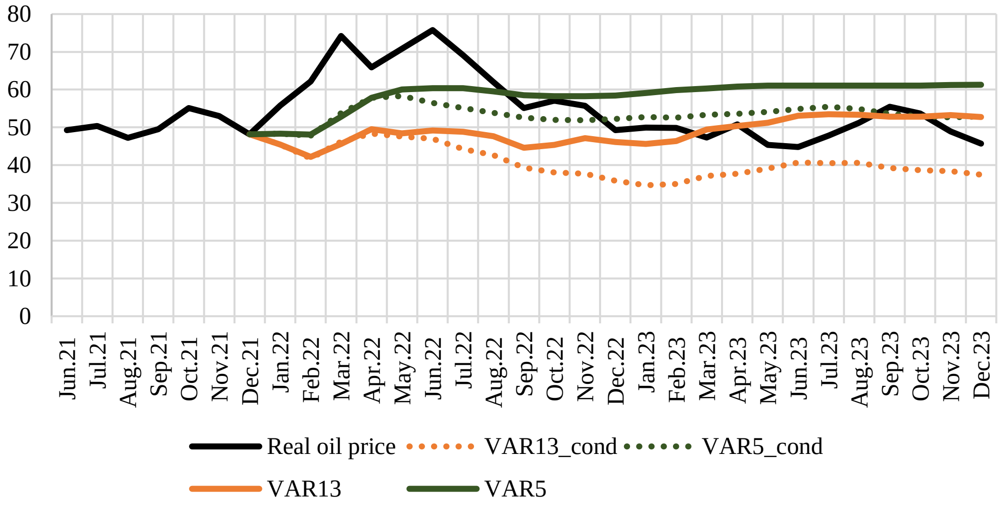
Conditional Forecasting

In order to improve the quality of vector autoregressive models, conditional forecasts were calculated. Conditional forecasting is a scenario-based forecast, where the future path of individual variables is assumed. This is different from an unconditional forecast, where no assumptions are made about this path.

The VAR5 and VAR13 models used EIA oil production and stocks forecasts. The conditional forecast was calculated on the 2003-2021 sample, and the out-of-sample forecast was made for 2022-2023.

The assessment results showed (Graph 3) that the conditional forecast for the VAR13 model did not show better results compared to the unconditional forecast for the same model (RMSE 16.7 versus RMSE 12.9, on average over the entire forecast horizon). At the same time, the conditional forecast for the VAR5 model was better than the original model (RMSE 9.1 versus RMSE 11.2).

Graph 3. Comparing Conditional and Unconditional Forecasts of Real Oil Price for 2022-2023



Source: EIA, authors' computations

To assess the robustness of the models, we decided to test the conditional forecasts for them on all rolling samples. However, the EIA provides forecasts each month that are limited to December of the following year (i.e., starting in January, each month the forecast values are reduced by one month, in December of the current year forecasts are available from December to December of the following year). In this regard, the forecast horizon was reduced to 13 months and only the VAR5 model was tested (Table 2). The conditional forecast for this model was improved on the horizon from the 6th to the 13th months.

Table 2. Root-Mean Squared Error (RMSE) of the Forecast for Conditional and Unconditional Forecasts Relative to RW

Model \ Months	1	2	3	4	5	6	7	8	9	10	11	12	13
VAR5	0.99	**0.98**	**0.88**	**0.81**	**0.75**	0.76	0.79	0.81	0.81	0.83	0.84	0.86	0.86
VAR5_cond	0.99	1.00	0.91	0.83	0.77	**0.75**	**0.75**	**0.76**	**0.76**	**0.77**	**0.77**	**0.78**	**0.79**

Note: In the table, a mean squared error value of less than one indicates that the given model outperforms the random walk model in the specified period. Values in bold indicate the superiority of the model compared to the other models presented.

Assessing the Forecasts Generated by the Models and Those of International Organizations

To assess the forecasting quality, the above models were compared with the actual dynamics of oil prices and the forecasts of Consensus Ecs. and EIA.

Given that Consensus Ecs. publishes quarterly forecasts, the monthly forecasts of the VAR5, VAR13 models as well as the estimates from the EIA and futures were converted to quarterly form (on average per quarter). In addition, given the limited forecast series from the EIA (mentioned in the previous section), the EIA estimates were compared on a horizon of up to 4 quarters.

The comparison of forecast estimates was carried out on the horizon of 8 and 4 quarters from Q1 2018 to Q4 2023 (a total of 24 sets of quarterly forecasts). The accuracy of the forecasts was assessed by the mean absolute percentage error, which is used to compare different models for one series (MAPE). In our case, the forecasts for the VAR5, VAR13 models are estimates of the real price of oil, and the forecasts of Consensus Ecs., EIA and futures are the nominal price of oil.

Thus, according to the results of comparative analysis, the forecasts of international organizations were on average inferior to the VAR5 and VAR13 models, except for the forecast for the first quarter (Table 3). When forecasting for one quarter, the Consensus Ecs. estimates were more accurate. In general, a smaller deviation from the fact on the forecast horizon from 2 to 4 quarters was observed for the VAR5 model, and from 5 to 8 quarters – for VAR13.

Table 3. Mean Absolute Percentage Error (MAPE)

Quarter	VAR5	VAR13	Consensus	Futures	EIA
1	9.8%	10.4%	**7.9%**	8.8%	8.9%
2	**16.5%**	17.9%	18.2%	18.7%	19.0%
3	**19.5%**	23.1%	23.5%	22.7%	23.4%
4	**24.4%**	24.5%	27.8%	28.3%	28.0%
5	28.6%	**22.8%**	29.4%	28.6%	-
6	31.8%	**19.2%**	30.7%	30.6%	-
7	36.0%	**19.9%**	32.8%	34.5%	-
8	35.5%	**18.8%**	34.8%	37.6%	-

Next, to test the forecasting qualities of the models, the same exercise was performed (results in Table 3), but in the VAR5 and VAR13 models, one of the dummy variables (March 2020, the COVID-19 pandemic) was replaced by January 2016 (US shale oil boom, weak global demand). As a result, the accuracy of the

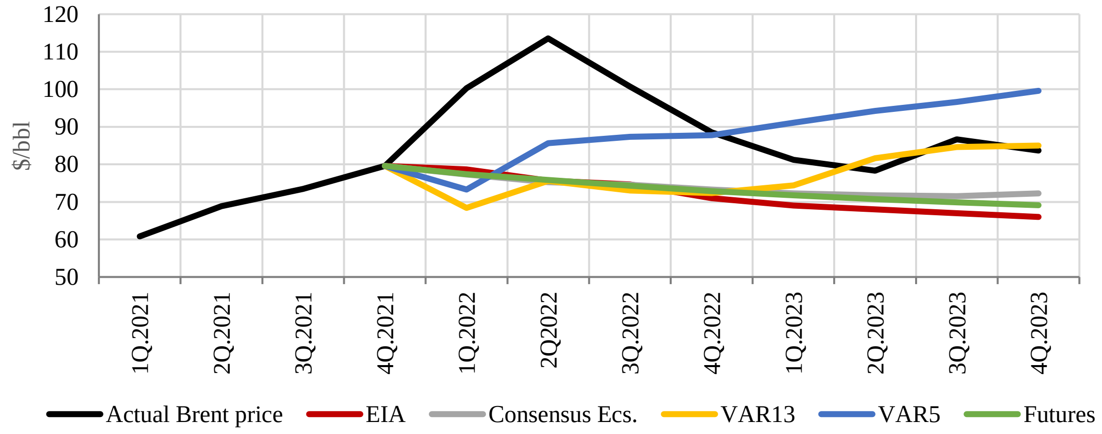
VAR5 forecasts deteriorated sharply over the entire forecast horizon. In this case, lower errors up to Q5 inclusive apply to the Consensus Ecs. and Futures forecasts. On a more medium-term horizon – to the VAR13 models (Table 4).

Table 4. Mean Absolute Percentage Error (MAPE)

Quarter	VAR5	VAR13	Consensus	Futures	EIA
1	10.8%	11.6%	**7.9%**	8.8%	8.9%
2	20.4%	22.5%	**18.2%**	18.7%	19.0%
3	24.7%	28.6%	23.5%	**22.7%**	23.4%
4	32.2%	31.7%	**27.8%**	28.3%	28.0%
5	35.5%	30.5%	29.4%	**28.6%**	
6	38.0%	**27.3%**	30.7%	30.6%	
7	41.0%	**25.2%**	32.8%	34.5%	
8	40.7%	**20.3%**	34.8%	37.6%	

Next, as an example, forecasts were made on the full sample (Graph 4). Consensus Ecs. forecasts for the specified period were taken from December 6, 2021, and the EIA – from January 11, 2022. For comparability, real oil price forecasts for the tested models (VAR5, VAR13) were converted to nominal values[7].

Graph 4. Comparing the Model-Generated Projections of Oil Price with Those of International Organizations



Source: EIA, Consensus Ecs., authors’ computations

As a result, the largest deviation from the actual price over the entire analyzed period was shown by the EIA and Consensus Ecs estimates. The tested VAR5 model captured the sharp price surge that occurred in Q2 2022 and almost accurately predicted the price in Q4 2022. The VAR13 model also captured the actual price increase to a greater extent in the medium term.

[7] Forecasts for inflation in the US are the Consensus Ecs.’s projections made on December 6, 2021

6. FORECAST COMBINING

Comparing Individual and Combined Models with Different Weight Diagrams

It is not recommended to build forecasts relying only on one model, even with minimal errors, since any model can show instability depending on the period. In addition, some models forecast better on shorter periods, and some on a longer horizon. In this case, a combination of forecasts can significantly improve the accuracy of forecasts.

In this regard, we decided to combine three models with the smallest errors and compare their results with individual models.

For the purposes of combining, we used several weighting methods:

1) equal weights (**Comb3_eqw**);

2) weighted average weights calculated as the ratio of the inversely proportional error value (in our case, RMSE) of forecasts for a single model to the sum of the inversely proportional error value (RMSE) of forecasts for all models (**Comb3_consw**). The weights themselves are constant, since they reflect the average error for each forecast horizon for all rolling samples. Thus, these weights are smoothed across the entire test sample.

$$\omega_{j,k} = \frac{1/RMSE_{j,k}}{\sum_{j=1}^{n} 1/RMSE_{j,k}} \qquad (2),$$

where

$w_{i,k}$ – forecast weights of model i, or a weight of each model

$RMSE_{j,k}$ – value of root-mean squared error (RMSE) of the out-of-sample forecasts of model i over the period $tk \in \{T0 + 1, \dots, T0 + K\}$.

$$Y^{f_i}_{combined} = \sum_{j=1}^{n} \omega_{j,k} * \hat{y}^{j}_{t_k} \qquad (3)$$

where, $Y^{f_i}_{combined}$ – a combined forecast of i-th indicator, $\hat{y}^{j}_{t_k}$ – the resulting out-of-sample forecasts of model j over the period $tk \in \{T0 + 1, \dots, T0 + K\}$.

A shortcoming of such approach is that it does not take into account the changes in quality of the model over time.

3) weighted average weights that change on each rolling sample (Comb3_rollw). For this purpose, a test section was allocated in each rolling period to obtain “pseudo-out-of-sample” forecasts. Then, the accuracy of forecasts of each model was assessed in this section. Based on these RMSE, weights were obtained for 24 months ahead for weighting the out-of-sample forecasts (inversely proportional to RMSE). As a result, 72 groups of weights for 24 months were calculated for all rolling samples.

The advantage of this approach is that it takes into account the change in the dynamics of indicators and the quality of models in each rolling sample, i.e. the most relevant weights are taken for forecasting.

Table 5 shows the forecast errors for all models. The exercise of combining the forecasts obtained by individual models did not change the picture significantly over the entire horizon. The models combined by all three weighting methods "won" over the individual models over the 2-month horizon, as well as from 8 to 13 months. In the remaining periods, the individual models retained the best forecasting properties: in the short term – ARMA and VAR5, and in the medium term – VAR13 (from 14 through 24 months).

Table 5. Comparison of errors between individual models and combined models weighted by different schemes

Month \ Model	ARMA (1,1)	VAR5	VAR13	Comb3_eqw	Comb3_consw	Comb3_rollw
1	**0.97**	0.99	1.20	1.00	0.99	1.00
2	0.99	0.98	1.10	0.97	0.97	**0.96**
3	1.00	**0.88**	1.00	0.90	0.89	0.93
4	0.99	**0.81**	0.94	0.84	0.83	0.87
5	0.98	**0.75**	0.90	0.79	0.78	0.80
6	0.98	**0.76**	0.93	0.79	0.78	0.81
7	0.98	**0.79**	0.96	0.80	0.80	0.83
8	0.96	0.81	0.96	**0.80**	**0.80**	0.81
9	0.96	0.81	0.94	0.80	**0.79**	0.80
10	0.95	0.83	0.91	**0.80**	**0.80**	**0.80**
11	0.95	0.84	0.87	**0.80**	**0.80**	0.81
12	0.95	0.86	0.85	0.81	0.81	**0.80**
13	0.92	0.86	0.80	0.80	0.79	**0.77**
15	0.93	0.89	**0.71**	0.77	0.76	0.77
18	0.93	0.92	**0.56**	0.73	0.68	0.66
21	0.92	0.93	**0.56**	0.73	0.68	0.65
24	0.92	0.91	**0.68**	0.76	0.73	0.73

Note: The table shows the root-mean squared errors of the models relative to the random walk model. A value less than one indicates that the given model outperforms the random walk model in the specified period. Values in bold indicate the superiority of the model compared to the other models presented.

Comparing Combined Forecasts of Different Models on Different Samples

In this section, we aim to test combined forecasts over different time periods and explore alternative methods of model combination.

Of all the presented combinations of models (Table 6), the best result on both samples is demonstrated by the combination of the RW, ARMA(1,1), VAR5 models. The combination of 3 models (ARMA(1,1),VAR5, VAR13) as presented in Table 4 has also shown good results.

Table 6. Comparing the Root-Mean Squared Error (RMSE) on Different Combinations, Relative to RW

Combination	Models	The 2003-2021 sample	The 2003-2016 sample
		The forecast for **2022-2023**	The forecast for **2017-2018**
2 models	**RW+ARMA(1,1)**	1.05	1.02
3 models	**2m+VAR5**	**0.74**	**0.55**
4 models	**3m+VAR13**	0.79	0.64
5 models	**4m+futures**	0.80	0.70
3 models*	**ARMA(1,1)+VAR5+VAR13**	0.77	0.60

Source: authors' computations

The combination of all 5 models on average over the forecast period (24 months) did not outperform the other models.

Next, we decided to test the selected two combinations of models at each horizon using a sliding window (Table 7). The results show that on average over all months, the combined model with VAR13 wins in forecast quality (from 3 through 24 months).

Table 7. Comparing Errors (RMSE) on Combined Models, Relative to RW, Rolling Window =180 Months

Month	**3 models***	**3 models**	**Month**	**3 models***	**3 models**
	ARMA(1,1) +VAR5 **+VAR13**	ARMA(1,1)+VAR5 **+ RW**		ARMA(1,1)+VAR5 **+VAR13**	ARMA(1,1)+VAR5 **+RW**
1	0.99	**0.95**	13	**0.79**	0.92
2	0.97	**0.96**	14	**0.78**	0.92
3	**0.89**	0.92	15	**0.76**	0.92
4	**0.83**	0.89	16	**0.72**	0.92
5	**0.78**	0.87	17	**0.70**	0.92
6	**0.78**	0.88	18	**0.68**	0.93
7	**0.80**	0.89	19	**0.68**	0.93
8	**0.80**	0.90	20	**0.66**	0.93
9	**0.79**	0.90	21	**0.68**	0.93
10	**0.80**	0.91	22	**0.71**	0.93
11	**0.80**	0.91	23	**0.73**	0.92
12	**0.81**	0.92	24	**0.73**	0.92

Source: authors' computations

7. CONCLUSION

Oil price dynamics play an important role in forecasting macroeconomic variables, both in terms of determining scenario-based development options and as an assumption for forecasting other variables (especially relevant for countries-exporters of primary goods). In this regard, oil price forecasting plays an important role in determining economic policy, including monetary policy.

At the same time, we recommend using the developed models as an additional tool for developing scenarios within forecast rounds, which will strengthen the analytical basis and efficiency.

Further work in this area can be based on the use of machine learning models (neural networks, etc.) and obtaining daily forecasts for oil prices. It is also possible to build models using financial indicators (interest rates, US dollar index, commodity indices, etc.).

8. LIST OF REFERENCES

1. Crude Oil Price Forecasting Using Hybrid Support Vector Machine, Lee Jo Xian, Shuhaida Ismail, Aida Mustapha, Mohd Helmy Abd Wahab, Syed Zulkarnain Syed Idrus, International Conference on Technology, Engineering and Sciences. 2020

2. Crude Oil Price Prediction using Artificial Neural Network. Nalini Gupta, Shobhit Nigam, The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40) April 6-9, 2020, Warsaw, Poland

3. Forecasting crude oil price using LSTM neural networks. Kexian Zhang* and Min Hong School of Economics and Management, Hunan Institute of Technology, Hengyang 421002, China, Data Science in Finance and Economics, Volume 2, Issue 3, 163–180, 07 July 2022

4. Forecasting Crude Oil Prices Using an ARIMA-ANN Hybrid Model. Hleil Alrweili and Haitham Fawzy, Journal of Statistics Applications & Probability 11, No. 3, 845-855 (2022), p.845

5. Benjamin Beckers, Samya Beidas-StromForecasting the Nominal Brent Oil Price with VARs—One Model Fits All? https://www.imf.org/external/pubs/ft/wp/2015/wp15251.pdf

6. Forecasting the Price of Oil, Handbook of Economic Forecasting (2013) https://doi.org/10.1016/B978-0-444-53683-9.00008-6

7. What central bankers need to know about forecasting oil prices (2014) https://doi.org/10.1111/iere.12074

8. A new hybrid model for forecasting Brent crude oil price (2020) https://doi.org/10.1016/j.energy.2020.117520

9. Kilian L., Zhou X. The Econometrics of Oil Market VAR Models. Federal Reserve Bank of Dallas. March 2020

10. Manescu C., Robays I.V. Forecasting the brent oil price addressing time-variation in forecast performance. Working paper series №1735.September 2014

11. EIA Short-Term Energy Outlook Crude Oil Price Forecasts

12. Ron Alquist, Lutz Kilian, and Robert J. Vigfusson. Forecasting the Price of Oil. Board of Governors of the Federal Reserve System. International Finance Discussion Papers 1022, July 2011

13. Baumeister C., Kilian L. Working Paper/Document de travail 2013-28 Forecasting the Real Price of Oil in a Changing World: A Forecast Combination Approach.

14. Baumeister C., Kilian L. and Thomas K. Lee. Are There Gains from Pooling Real Time Oil Price Forecasts? Bank of Canada Working Paper 2014-46 October 2014