430

JOURNAL
OF EUROPEAN ECONOMY
Vol. 22. № 3 (86). July–September 2023
Publication of West Ukrainian National University

Primary Sector Economics

Vasileios ANASTASIADIS,
Evangelos SISKOS

TIME SERIES ANALYSIS FOR FORECASTING CRUDE OIL PRICES

Abstract

Many analysts, policymakers, and researchers have grown increasingly concerned about the fluctuation of international crude oil prices. That is because oil prices reflect many macroeconomic and financial indicators (GDP, unemployment, inflation, S&P 500 Index, Nasdaq Composite Index), and conditions in a variety of financial and goods markets. This paper highlights the most appropriate model for estimating and forecasting West Texas Intermediate (WTI) crude oil monthly prices by comparing three hybrid models – ARMA-GARCH, ARMA-EGARCH, and ARMA-FIGARCH. Finally, among these models, the paper considers that the ARMA-EGARCH(1,20) model emerges as the most efficacious model for the prediction of West Texas Intermediate (WTI) crude oil monthly price returns.

Anastasiadis, Vasileios, M.Sc. in Oil and Gas Management and Transportation, University of Western Macedonia, Kozani, Greece. ORCID: 0009-0004-8658-5798 Email: billys25.1996@gmail.com
Siskos, Evangelos, DSc (Econ), Professor of International, European and Black Sea Economic Relations, Department of International and European Economic Studies, University of Western Macedonia, Kozani, Greece. ORCID: 0000-0002-5221-4444 Email: esiskos@uowm.gr

Key Words:

international crude oil prices; forecasting; ARMA; GARCH; returns; Eviews.

JEL: Q47.

7 figures, 6 tables, 5 formulae, 14 references.

Problem Statement

This paper aims to identify the most suitable model for estimating and forecasting the returns for monthly prices of West Texas Intermediate (WTI) crude oil. Specifically, the study focuses on the period from March 2022 to May 2022. The three hybrid models compared in this analysis are ARMA-GARCH, ARMA-EGARCH, and ARMA-FIGARCH, with the goal of determining the model that returns the most accurate forecasts.

The analysis is carried out using the Box-Jenkins methodology and GARCH models, including ARMA-GARCH(1,2), ARMA-GARCH(1,20), ARMA-EGARCH(1,2), ARMA-EGARCH(1,20), ARMA-FIGARCH(1,2), and ARMA-FIGARCH(1,20). Among these models, the paper identifies the most suitable hybrid model based on its ability to minimize Akaike and Schwarz criteria.

This research aims to provide insights into the forecast accuracy of the selected model by employing forecast rating indices, such as root mean squared error (RMSE), absolute mean error (MAE), and absolute average error rate (MAPE). The findings of this study will contribute to the understanding of which hybrid model is most effective for forecasting West Texas Intermediate (WTI) crude oil prices return.

Literature Review

Moosa and Al-Loughani (1994) examined the relationship between spot and future crude oil prices (WTI) with various tests and with the GARCH-M model (1,1). They used monthly data from 1/1986 to 7/1990 for three-time series. Initially, for spot price, secondly for futures price for three months later (f_3) and six months later (f_6). The survey results showed that futures contracts could not be unbiased and efficient for forecasting spot prices. Finally, they concluded that the GARCH-M process can decently model the difference between risk premiums over time but cannot be generalized.

Huang et al. (1996) used the vector autoregressive (VAR) model to examine the relationship between future daily crude oil WTI returns and U.S. stock returns. They found that crude oil returns (WTI) do not significantly affect noteworthy stock market indexes, such as the S&P 500. The analysis was based on daily closing prices for forward contracts in the near future on the New York Stock Exchange (NYMEX) for the period 11/05/1983 – 16/03/1990.

Sadorsky (1999) combined VAR and GARCH models to show that oil price and oil price volatility play a vital role in the economy and stock exchange yields, while changes in economic activity have a low impact on oil prices. The seasonally adjusted industrial production index (1982 = 100), the level of three-month government bonds, the seasonally adjusted producer price index (1982 = 100), the seasonally adjusted index of 500 ordinary shares (1967 = 100), the seasonally adjusted consumer price index (1982 = 100), oil prices, and share yields were among the used variables.

Radchenko (2005) applied a variable voltage model and a back-to-back procedure with error conditions that do not have the characteristics of the white noise model. He used weekly data from oil and gasoline prices in the United States from 3/1991 to 2/2003. The differences in their logarithmic values were also used, as there was low inflation (1.54 – 3.58%) during the considered period. In addition, in October 1983 there was a considerable tax increase on petrol prices. For this reason, the researcher added a pseudo variable, equal to 0, for the period before October 1993, whereas otherwise, it would be equal to 1. Finally, the investigator split the entire sample into two models (low and high variability).

Sadorsky and Basher (2006) examined the prices of heating oil, unleaded petrol, natural gas, and futures prices for crude oil (WTI) traded on the New York Mercantile Exchange (NYMEX). For crude oil, they used the data from 5/2/1988 – 31/1/2003 and compared the GARCH (1,1), TGARCH (1,1), BIGARCH (1,1) models, various models of moving medium, linear regression, and exponential smoothing. They concluded that GARCH (1,1) predicts crude oil price volatility better.

Hung et al. (2008) investigated the efficacy of the day-ahead Value at Risk (VaR), with three different distributions (normal distribution, *t*-student distribution, distribution with heavy tails), in the GARCH (1,1) (GARCH-N, GARCH-*t*, GARCH-HT). Specifically, they used WTI prices (for the period 09/09/1996 – 31/08/2006) and BRENT prices (for the period 05/11/1996 – 31/08/2006). The results showed that the heavy-tailed distribution gives more satisfactory results for calculating Value at Risk.

Muradov et al. (2018) developed an econometric model that calculated the average prices of WTI and BRENT, respectively. They used annual data from 1975 to 2017. In fact, they introduced pseudo-variables from 2008 to 2015 due to the crises and the sharp fall in oil prices. In addition, the same model used the variable "@TREND" as the voltage average. An ARIMA model was included with the first and ninth hysteresis of the autoregressive model and the tenth hysteresis of the moving average. The model was found to be adequate and useful for forecasting.

Yang et al. (2002) explored the short- and long-term relationship between oil demand, GDP, oil, gas and coal prices with an ECM (Error Correction Model). Thet tested various scenarios which assumed a 4% decrease in the production of OPEC in order to calculate the elasticity of demand. As a result, oil prices increased but could also decrease (if there was a recession). Remarkable is their hypothesis that a decrease in oil production by OPEC may cause stagflation and eventually lower oil prices (Organization of the Petroleum Exporting Countries, 2022).

Mohammadi and Su (2010) examined the effectiveness of various ARIMA and GARCH models in predicting average and weekly crude oil price volatility for 01/2009 – 10/2009. They used data from 11 countries, both importers and exporters (Algeria, Canada, China, Dubai, Indonesia, Norway, and Russia), for the period 03/01/1997 – 13/02/2009 and concluded that the APARCH model is preferred. A neural network (ANN) showed that the oil futures market is an inefficient market that ensures profitable transactions. In addition, it exhibited the predictability of oil prices relative to other models, for example, buy and hold rule (BH; assumes prices are rising permanently), technical analysis of conventional moving average, random walk, and zero risk interest rate (90 days government bond, T-bil). In particular, data from 1985 to 2007 were used for oil supply, crude oil distillation capacity, petroleum consumption of countries other than OECD, the capacity of U.S. refineries and overcapacity.

Mirmirani and Li (2004) compared VAR and neural network (ANN) techniques for predicting crude oil prices in the United States. In particular, historical oil prices, previous supply, previous energy consumption and money supply (M1) were used. They included monthly data for 01/1980 – 12/2002 on light sweet crude oil prices. Finally, the results showed that the ANN model showed better results than the VAR model.

Yu et al. (2017) showed that the support vector machine (SVM) application exhibited better results than the ANN, ARIMA, and ARFIMA models for BRENT and WTI values.

Wei, Wang, Li, and Chen (2022) conducted a study to assess the effects of the pandemic on the long-term volatility and correlation of gold and crude oil prices. Initially, they employed a DCC GARCH model (Dynamic Conditional Correlation GARCH) to quantify the pandemic's impact on these commodities' volatility. Subsequently, they utilized the DCC-MIDAS GARCH model (Dynamic Conditional Correlation – Mixed Data Sampling GARCH) to assess the influence of the pandemic on the long-term correlation between gold and crude oil markets. The survey data was divided into two categories. The first category consisted of London-based daily settlement prices for gold and BRENT crude oil. The second category included monthly data on infections and their volatility. The findings indicated a significant positive effect of the pandemic on the long-term volatility of gold and crude oil prices. Moreover, as the pandemic duration increased, these effects became more pronounced. Additionally, the pandemic positively impacted the long-term correlation between the two markets.

Methodology and Empirical Analysis

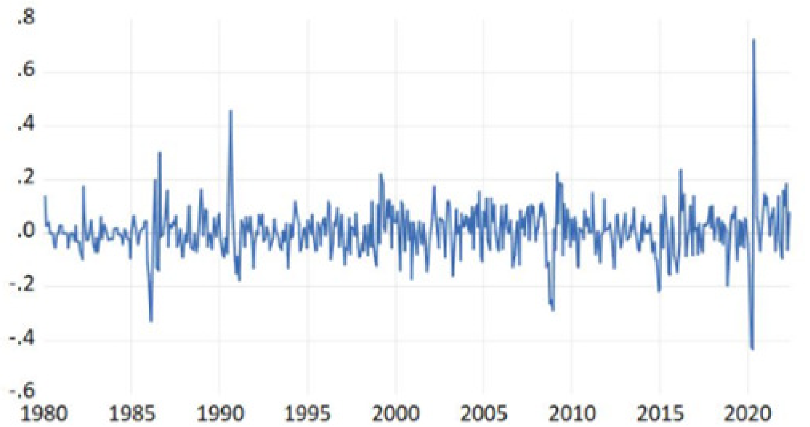
This paper used monthly WTI crude oil prices from January 1980 to May 2022. The data were extracted from the E.I.A. database, the U.S. Federal Bureau of Energy (U.S. Energy Information Administration, 2023). The data were divided into two groups. The first data group (January 1980 – February 2022) was used to train the model, while the second group (March 2022 – May 2022) helped examine the model's predictive capacity.

This paper aimed to identify the most suitable model for estimating and forecasting the returns for monthly prices of West Texas Intermediate (WTI) crude oil. The analysis was carried out using the statistical software EViews 12.0.

As seen from Figure 1, there were no clear outliers in WTI crude oil price returns (except for 2020), nor were there any particular trends.

Figure 1

Crude oil price returns



At this stage, the purpose was to determine whether the time series is stationary or if the first differences need to be calculated. The Dickey-Fuller, Phillips-Peron, and KPSS tests were all used to test for stationarity. The results of these tests suggest that the time series is indeed stationary. Specifically, in the augmented Dickey-Fuller test, the value of probability (prob.) is 0. This test determined the maximum number of lags based on the Akaike Information Criterion. Additionally, the Phillips-Peron test is equal to 0, and the LM-Stat in the KPSS test was statistically significant at all three significance levels. Subsequently, an ARMA(p,q) model is examined using different statistically significant lag values. Appendix A summarises the number of statistically significant lags, the amount of SIGMASQ, the adjusted R^2 and the value of the Akaike and the Schwarz test. The ARMA(1,2) model is advantageous because of its simplicity, while the ARMA(1,20) model can be beneficial when considering the potential presence of seasonality. Tables 1 and 2 show the estimate of the ARMA(1,2) and ARMA(1,20), respectively, as presented in Eviews.

Table 1

Estimation of ARMA(1,2)

Dependent Variable: RETURN_WTISPLC
Method: ARMA Maximum Likelihood (OPG – BHHH)
Date: 08/26/22 Time: 01:50
Sample: 1980M02 2022M02
Included observations: 505
Convergence achieved after 92 iterations
Coefficient covariance computed using outer product of gradients

Variance	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006373	0.005401	1.179999	0.2386
AR(1)	0.275297	0.031925	8.623283	0.0000
MA(2)	-0.37443	0.041643	-3.300509	0.0010
SIGMASQ	0.007937	0.000257	30.91357	0.0000
R-squared	0.076006	Mean dependent var		0.006336
Adjusted R-squared	0.070473	S.D. dependent var		0.092774
S.E. of regression	0.089446	Akaike info criterion		-1.982293
Sum squared resid	4.008256	Schwarz criterion		-1.948832
Log likelihood	504.5291	Hannan-Quinn criter.		-1.969169
F-statistic	13.73705	Durbin-Watson stat.		1.991886
Prob(F-statistic)	0.000000			
Inverted AR Roots	.28			
Inverted MA Roots	.37	-.37		

Initially, after the possible models have been selected, it is necessary to check if the models represent autocorrelation. For this particular purpose, the Ljung–Box test is used:

$$Q = n(n+2)\sum_{k=1}^{h}\frac{\hat{\rho}_k^2}{n-k} \quad (1)$$

where *n* is the sample size; *h* is the number of autocorrelation coefficients; $\hat{\rho}_k$ refers to the size of autocorrelation at lag *k*.

Table 2

Estimation of ARMA(1,20)

Dependent Variable: RETURN_WTISPLC
Method: ARMA Maximum Likelihood (OPG – BHHH)
Date: 08/26/22 Time: 01:49
Sample: 1980M02 2022M02
Included observations: 505
Convergence achieved after 60 iterations
Coefficient covariance computed using outer product of gradients

Variance	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006676	0.006776	0.985277	0.3250
AR(1)	0.251944	0.024376	10.33591	0.0000
MA(20)	0.117502	0.050181	2.341548	0.0196
SIGMASQ	0.007935	0.000219	36.30569	0.0000
R-squared	0.076254	Mean dependent var		0.006336
Adjusted R-squared	0.070722	S.D. dependent var		0.092774
S.E. of regression	0.089434	Akaike info criterion		-1.982071
Sum squared resid	4.007181	Schwarz criterion		-1.948610
Log likelihood	504.4730	Hannan-Quinn criter.		-1.968947
F-statistic	13.78553	Durbin-Watson stat.		1.937210
Prob(F-statistic)	0.000000			
Inverted AR Roots	.25			
Inverted MA Roots	.89-.14i	.89+.14i	.80+.41i	.80-.41i
	.64+.64i	.64-.64i	.41+.80i	.41-.80i
	.14+.89i	.14-.89i	-.14+.89i	-.14-.89i
	-.41-.80i	-.41+.80i	-.64-.64i	-.64-.64i
	-.80-.41i	-.80+.41i	-.89+.14i	-.89-.14i

The hypotheses of the test are:

- H_0: No autocorrelation;
- H_1: There is autocorrelation, the data are not distributed independently and they show a serial correlation.

The correlogram of residuals makes it apparent that most prob values of the Ljung–Box test are more significant than 0.5. This means that the hypothesis of no autocorrelation cannot be rejected (Figures 2 and 3).

Figure 2

Correlogram of residuals, ARMA(1,2)

Correlogram of Residuals

Date: 06/24/22 Time: 17:45
Sample (adjusted): 1980M02 2022M02
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.002	0.002	0.0017	
		2	0.014	0.014	0.0956	
		3	-0.053	-0.053	1.5335	0.216
		4	-0.091	-0.092	5.8023	0.055
		5	-0.018	-0.016	5.9607	0.114
		6	-0.052	-0.053	7.3354	0.119
		7	0.005	-0.004	7.3497	0.196
		8	0.000	-0.009	7.3497	0.290
		9	-0.000	-0.009	7.3497	0.393
		10	0.025	0.015	7.6626	0.467
		11	0.012	0.010	7.7406	0.560
		12	-0.017	-0.022	7.8966	0.639
		13	-0.029	-0.029	8.3390	0.683
		14	-0.013	-0.009	8.4290	0.751
		15	-0.058	-0.058	10.157	0.681
		16	-0.009	-0.014	10.196	0.748
		17	0.026	0.021	10.539	0.785
		18	-0.066	-0.079	12.844	0.684
		19	-0.012	-0.029	12.919	0.742
		20	0.113	0.114	19.660	0.352
		21	0.010	0.000	19.715	0.412
		22	0.023	0.004	19.988	0.459

Figure 3

Correlogram of residuals, ARMA(1,20)

Correlogram of Residuals

Date: 06/24/22 Time: 17:49
Sample (adjusted): 1980M02 2022M02
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.029	0.029	0.4335	
		2	-0.099	-0.100	5.4308	
		3	-0.053	-0.047	6.8520	0.009
		4	-0.086	-0.094	10.634	0.005
		5	-0.010	-0.016	10.688	0.014
		6	-0.034	-0.056	11.284	0.024
		7	0.013	0.003	11.370	0.045
		8	0.002	-0.018	11.373	0.078
		9	0.000	-0.005	11.373	0.123
		10	0.022	0.013	11.618	0.169
		11	0.014	0.012	11.715	0.230
		12	-0.015	-0.015	11.831	0.297
		13	-0.026	-0.021	12.189	0.350
		14	-0.007	-0.005	12.213	0.429
		15	-0.057	-0.062	13.897	0.381
		16	0.013	0.011	13.983	0.451
		17	0.035	0.017	14.634	0.478
		18	-0.068	-0.078	17.063	0.382
		19	-0.012	-0.015	17.134	0.445
		20	0.005	-0.008	17.146	0.513
		21	0.012	0.001	17.228	0.574
		22	0.027	0.015	17.627	0.612

Furthermore, the correlogram of residuals squared shows that apparently almost all lags are not statistically significant, which means that all information has been included from the model (Figures 4 and 5).

Figure 4

Correlogram of residuals squared, ARMA(1,2)

Correlogram of Residuals Squared

Date: 06/24/22 Time: 17:48
Sample (adjusted): 1980M02 2022M02
Included observations: 505 after adjustments

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.195	0.195	19.371	0.000
		2	0.213	0.182	42.405	0.000
		3	0.017	-0.056	42.559	0.000
		4	0.003	-0.032	42.564	0.000
		5	0.000	0.015	42.565	0.000
		6	0.014	0.020	42.662	0.000
		7	0.020	0.013	42.872	0.000
		8	-0.005	-0.019	42.885	0.000
		9	0.004	0.002	42.895	0.000
		10	-0.014	-0.010	42.999	0.000
		11	-0.012	-0.010	43.078	0.000
		12	-0.016	-0.009	43.212	0.000
		13	-0.013	-0.006	43.302	0.000
		14	-0.024	-0.018	43.603	0.000
		15	-0.014	-0.004	43.700	0.000
		16	-0.010	0.001	43.748	0.000
		17	0.004	0.009	43.756	0.000
		18	0.023	0.024	44.046	0.001
		19	-0.021	-0.034	44.274	0.001
		20	0.021	0.023	44.512	0.001
		21	-0.016	-0.011	44.650	0.002
		22	-0.006	-0.012	44.671	0.003

Figure 5

Correlogram of residuals squared, ARMA(1,20)

Correlogram of Residuals

Date: 06/24/22 Time: 17:49
Sample (adjusted): 1980M02 2022M02
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.029	0.029	0.4335	
		2	-0.099	-0.100	5.4308	
		3	-0.053	-0.047	6.8520	0.009
		4	-0.086	-0.094	10.634	0.005
		5	-0.010	-0.016	10.688	0.014
		6	-0.034	-0.056	11.284	0.024
		7	0.013	0.003	11.370	0.045
		8	0.002	-0.018	11.373	0.078
		9	0.000	-0.005	11.373	0.123
		10	0.022	0.013	11.618	0.169
		11	0.014	0.012	11.715	0.230
		12	-0.015	-0.015	11.831	0.297
		13	-0.026	-0.021	12.189	0.350
		14	-0.007	-0.005	12.213	0.429
		15	-0.057	-0.062	13.897	0.381
		16	0.013	0.011	13.983	0.451
		17	0.035	0.017	14.634	0.478
		18	-0.068	-0.078	17.063	0.382
		19	-0.012	-0.015	17.134	0.445
		20	0.005	-0.008	17.146	0.513
		21	0.012	0.001	17.228	0.574
		22	0.027	0.015	17.627	0.612

Consequently, the models are tested for heteroscedasticity with the ARCH test. Specifically, if the b_1 coefficient is statistically significant, there is heteroscedasticity in the residuals.

$$\hat{u}_t^2 = b_0 + b_1 \hat{u}_{t-1}^2 + \varepsilon_t \quad (2)$$

H_0: data is homoscedastic

H_1: data is heteroscedastic

Table 3 shows the ARCH test for ARMA(1,2) and the Table 4 shows the same for ARMA(1,20).

Table 3

ARCH test for ARMA(1,2)

Heteroskedasticity Test: ARCH				
F-statistic	19.90549	Prob. F (1,502)		0.0000
Obs*R-squared	19.22258	Prob. Chi-Square (1)		0.0000
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 06/24/22 Time: 19:02				
Sample (adjusted): 1980M03 2022M02				
Included observations: 504 after adjustments				
Variance	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006368	0.001408	4.522461	0.0000
RESId^2(-1)	0.195289	0.043771	4.461557	0.0000
R-squared	0.038140	Mean dependent var		0.007921
Adjusted R-squared	0.036224	S.D. dependent var		0.031203
S.E. of regression	0.030633	Akaike info criterion		-4.129513
Sum squared resid	0.471069	Schwarz criterion		-4.112757
Log likelihood	1042.637	Hannan-Quinn criter.		-4.122940
F-statistic	19.90549	Durbin-Watson stat.		2.070133
Prob(F-statistic)	0.000010			

Table 4

ARCH test for ARMA(1,20)

Heteroskedasticity Test: ARCH				
F-statistic	17.48205	Prob. F (1,502)		0.0000
Obs*R-squared	16.96104	Prob. Chi-Square (1)		0.0000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 06/24/22 Time: 19:03
Sample (adjusted): 1980M03 2022M02
Included observations: 504 after adjustments

Variance	Coefficient	Std. Error	t-Statistic	Prob.
C	0.006461	0.001514	4.266360	0.0000
RESId^2(-1)	0.183443	0.043874	4.181154	0.0000
R-squared	0.033653	Mean dependent var		0.007919
Adjusted R-squared	0.031728	S.D. dependent var		0.033622
S.E. of regression	0.033084	Akaike info criterion		-3.975552
Sum squared resid	0.549476	Schwarz criterion		-3.958796
Log likelihood	1003.839	Hannan-Quinn criter.		-3.968980
F-statistic	17.48205	Durbin-Watson stat.		2.061614
Prob(F-statistic)	0.000034			

The value of prob Chi-Square and the coefficient are statistically significant, so the null hypothesis of homoscedasticity cannot be accepted. Thus, the different types of Generalized AutoRegressive Conditional Heteroskedasticity model, GARCH(p,q), is used in order to fix the problem of heteroscedasticity, since the GARCH model aims to fix the problem by estimating the conditional variance through the equation:

$$GARCH\text{: } \sigma_t^2 = \alpha_0 + \sum_{p=1}^{\rho} \alpha_p \, \varepsilon_{t-p}^2 + \sum_{q=1}^{q} \beta_q \sigma_{t-q}^2 \quad (3)$$

where *e* is the residuals, the σ^2 is the variance, and the *p*, *q* are the numbers of lags.

According to Hamilton (1994), negative shocks impact variability more than positive shocks. Therefore, it is important to consider the sign of change alongside its magnitude in view of the leverage effect. Nelson (1991) proposed a solution by expressing the variance in logarithmic form and introducing a fourth term in the equation that represents the sign of the error. This approach allows for asymmetric variability between positive and negative returns.

$$EGARCH\ log(\sigma^2_t) = \alpha_0 + \sum_{j=1}^{\rho}\left(\gamma_t \frac{\varepsilon_{t-j}}{\sigma_{t-j}} + \alpha_\rho \frac{|\varepsilon_{t-j}|}{|\sigma_{t-j}|}\right) + \sum_{i=1}^{q} \beta_i log\left(\sigma^2_{t-i}\right) \quad (4)$$

- If $\gamma = 0$: A positive shock has the same effect as a negative shock;
- If $\gamma < 0$: A positive shock decreases volatility, a negative shock increases it;
- If $\gamma > 0$: A positive shock increases the volatility, a negative shock decreases it.

FIGARCH models are primarily needed for time series with a strong memory, where a shock affects the variability for multiple lags. Ding, Granger and Engle (1993) introduced the model FIGARCH(1,d,1):

$$FIGARCH(1,d,1): \sigma_t = \omega + \beta_1\sigma_{t-1} + [1 - \beta_1 * L - (1 - \varphi_1 * L) * (1 - L)^d] * \varepsilon_t^2 \quad (5)$$

where: σ_t is conditional dependent standard deviation, ω is constant term, L is lag operator, φ is density function of a normal probability distribution, d is long-memory parameter.

When $0 < d < 1$, the series is stationary with solid memory. Conversely, if $d = 1$, the process has a unit root, and if $d = 0$, the model essentially transforms into a simple GARCH(1,1).

Finally, the ARMA model (1,20) with the EGARCH method is selected because it minimizes the AIC and the SIC tests (Appendix B). Table 5 shows the model estimation as shown in Eviews.

Similarly, an LM GARCH model is reevaluated, but the value of Prob and the coefficient are not significant. As a result, the hypothesis of homoscedasticity cannot be rejected. Considering the above findings, the model can be utilized for prediction (Table 6).

The EGARCH model does not exhibit any serial correlation. Figure 6 shows the correlogram of standardized residuals for the Prob values.

Table 5

Model estimation ARMA – EGARCH (1,20)

Dependent Variable: RETURN_WTISPLC
Method: ML ARCH – Student's t distribution (OPG – BHHH / Marquardt steps)
Date: 08/26/22 Time: 02:00
Sample (adjusted): 1980M03 2022M02
Included observations: 504 after adjustments
Convergence achieved after 49 iterations
Coefficient covariance computed using outer product of gradients
MA Backcast: 1978M07 1980M02
Presample variance: backcast (parameter = 0.7)
LOG(GARCH) = C(4) +C(5)*ABS(RESID(-1)/@SQRT(GARCH(-1))) + C(6)
*RESID(-1)/@SQRT(GARCH(-1)) + C(7)*LOG(GARCH(-1))

Variance	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.001802	0.004112	-0.438350	0.6611
AR(1)	0.246458	0.044050	5.594992	0.0000
MA(20)	0.122109	0.031452	3.882406	0.0001
Variance Equation				
C(4)	-0.732201	0.207194	-3.533886	0.0004
C(5)	0.374084	0.084501	4.426951	0.0000
C(6)	-0.187680	0.048720	-3.852189	0.0001
C(7)	0.913612	0.031891	28.64779	0.0000
T-DIST. DOF	9.969953	3.224968	3.091490	0.0020
R-squared	0.073209	Mean dependent var		0.006074
Adjusted R-squared	0.069510	S.D. dependent var		0.092679
S.E. of regression	0.089400	Akaike info criterion		-2.381210
Sum squared resid	4.004174	Schwarz criterion		-2.314185
Log likelihood	608.0650	Hannan-Quinn criter.		-2.354919
Durbin-Watson stat.	1.923282			
Inverted AR Roots	.25			
Inverted MA Roots	.89-.14i	.89+.14i	.80+.41i	.80-.41i
	.64+.64i	.64-.64i	.41+.80i	.41-.80i
	.14+.89i	.14-.89i	-.14+.89i	-.14-.89i

Figure 6

Correlogram of standardized residuals for the EGARCH model

Correlogram of Standardized Residuals

Date: 06/24/22 Time: 20:49
Sample (adjusted): 1980M03 2022M02
Q-statistic probabilities adjusted for 2 ARMA terms

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob*
		1	0.047	0.047	1.1136	
		2	-0.070	-0.073	3.6142	
		3	-0.041	-0.034	4.4698	0.034
		4	-0.057	-0.059	6.1270	0.047
		5	-0.032	-0.032	6.6397	0.084
		6	-0.023	-0.030	6.9138	0.141
		7	0.006	-0.000	6.9342	0.226
		8	-0.006	-0.016	6.9523	0.325
		9	0.029	0.025	7.3832	0.390
		10	0.002	-0.006	7.3853	0.496
		11	0.047	0.050	8.5157	0.483
		12	0.010	0.006	8.5667	0.574
		13	-0.060	-0.052	10.438	0.491
		14	-0.012	-0.002	10.515	0.571
		15	-0.084	-0.086	14.166	0.362
		16	-0.037	-0.032	14.871	0.387
		17	0.045	0.033	15.953	0.385
		18	-0.031	-0.052	16.453	0.422
		19	-0.001	-0.005	16.453	0.492
		20	-0.026	-0.041	16.817	0.536
		21	0.039	0.035	17.604	0.549
		22	0.030	0.020	18.079	0.582

ISSN 2519-4070

Table 6

Heteroscedasticity check (LM GARCH)

Heteroskedasticity Test: ARCH

F-statistic	0.010699	Prob. F (1,501)	0.9177
Obs*R-squared	0.010742	Prob. Chi-Square (1)	0.9175

Test Equation:
Dependent Variable: WGT_RESID^2
Method: Least Squares
Date: 06/24/22 Time: 20:55
Sample (adjusted): 1980M04 2022M02
Included observations: 503 after adjustments

Variance	Coefficient	Std. Error	t-Statistic	Prob.
C	1.004295	0.096994	10.35416	0.0000
WGT_RESId^2(-1)	0.004621	0.044670	0.103437	0.9177

R-squared	0.000021	Mean dependent var	1.008954
Adjusted R-squared	-0.001975	S.D. dependent var	1.924678
S.E. of regression	1.926577	Akaike info criterion	4.153335
Sum squared resid	1859.562	Schwarz criterion	4.170117
Log likelihood	-1042.564	Hannan-Quinn criter.	4.159919
F-statistic	0.010699	Durbin-Watson stat.	1.998744
Prob(F-statistic)	0.917657		

Discussion of Research Results

The results of this study suggest that the ARMA-EGARCH(1,20) model is the most effective model for forecasting returns for international crude oil prices (WTI) over the period of March to May 2022. Figure 7 shows the forecast of returns for crude oil prices (WTI) with the static model, while Figure 8 shows the dynamic model. The green line represents the actual values of WTI yields, and the blue line represents the predictions. The red dotted lines represent the confidence interval limits ($\pm$)2. The values of the indexes (root-mean-square error RMSE = $\sqrt{\frac{1}{T}\sum_1^T(\hat{Y}_t - Y_t)^2}$= 0,127), absolute mean error MAE = $\frac{1}{T}\sum_1^T\left|(\hat{Y}_t - Y_t)\right|$ = 0,124), mean absolute percentage error MAPE = $\frac{1}{T}\sum_1^T\left|\frac{\hat{Y}_t - Y_t}{Y_t}\right|$ = 133,41) are low, which means that the model has a relatively low error rate, and is a good fit to the data. Finally, appendix C shows the actual values of the returns against the forecasts in Eviews.

Figure 6

Static model for predicting oil price returns (WTI)

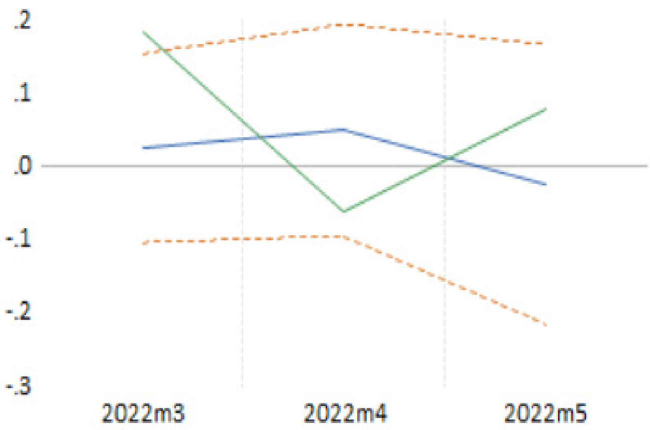
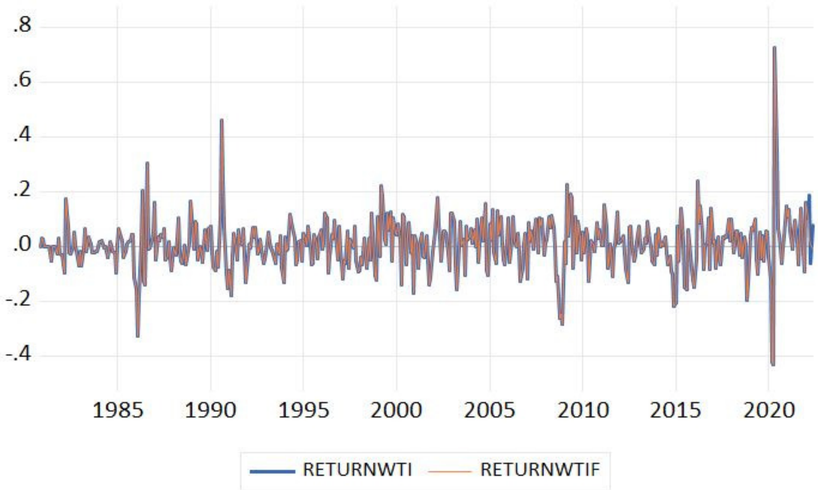


Figure 7

Dynamic model for predicting oil price yields (WTI)



The differences between the actual and predicted values were small for all three months (-0.16, 0.07, and -0.085 for March, April, and May, respectively). This suggests that the model is suitable for forecasting WTI prices. In fact, this further supports the conclusion that the model is effective.

Conclusions

Crude oil significantly impacts global economic and social development, so its price formation has attracted the attention of many governments, investors, analysts, and scholars. However, predicting the price of oil is a challenging process.

This paper aimed to determine the model that showed the most accurate forecast returns for international crude oil prices from March 2022 to May 2022. In particular, the data was collected from the U.S. Federal Bureau of Energy. (U.S. EIA, 2023). The analysis was based on the Box – Jenkins methodology and the GARCH (a generalized form of autoregressive bound heteroscedasticity) models and was carried out with the use of the Eviews 12.0 statistical program. Several models were tested, namely ARMA-GARCH(1,2), ARMA-GARCH(1,20), ARMA-EGARCH(1,2), ARMA-EGARCH(1,20), ARMA-FIGARCH(1,2), and ARMA-FIGARCH(1,20). ARMA – EGARCH(1,20) was the hybrid model selected in the end, as it minimizes Akaike and Schwarz checks.

Furthermore, the paper calculated the forecast rating indices of the root mean squared error (RMSE 0,127), the absolute mean error (MAE 0,124), and the absolute average error rate (MAPE 133,41). The above values are considered relatively low, which means that the error of the forecast is also relatively low, and the model is suitable for forecasting.

The final results of the forecasts against the actual values were calculated as follows

- for March, 0.023 against 0.183,
- for April, 0.009 against -0.061,
- for May, -0.009 against 0.076.

The model’s predictions were accurate for all three months, with a relatively small error. This suggests that the model is effective for forecasting WTI prices.

Additionally, it is clear that future research would benefit from using neural networks or machine learning algorithms, as in recent years, they have shown better results than traditional econometric methods.

References

Basher, S. A., & Sadorsky, P. (2006). Oil price risk and emerging stock markets. *Global finance journal*, *17*(2), 224-251. https://doi.org/10.1016/j.gfj.2006.04.001

Huang, R. D., Masulis, R. W., & Stoll, H. R. (1996). Energy shocks and financial markets. *Journal of Futures markets*, *16*(1), 1-27. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=900741

Hung, J. C., Lee, M. C., & Liu, H. C. (2008). Estimation of value-at-risk for energy commodities via fat-tailed GARCH models. *Energy Economics*, *30*(3), 1173-1191. https://doi.org/10.1016/j.eneco.2007.11.004

Mirmirani, S., & Cheng Li, H. (2004). A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. In J.M. Binner, G. Kendall & S.-H. Chen (Eds.), *Applications of Artificial Intelligence in Finance and Economics* (Advances in Econometrics, Vol. 19, pp. 203-223). Emerald Group Publishing Limited. https://doi.org/10.1016/S0731-9053(04)19008-7

Mohammadi, H., & Su, L. (2010). International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. *Energy Economics*, *32*(5), 1001-1008. https://doi.org/10.1016/j.eneco.2010.04.009

Moosa, I. A., & Al-Loughani, N. E. (1994). Unbiasedness and time varying risk premia in the crude oil futures market. *Energy Economics*, *16*(2), 99-105. https://doi.org/10.1016/0140-9883(94)90003-5

Muradov, A., Hasanli, Y., & Hajiyev, N. (2018). Crude Oil Price Forecasting Techniques in the World Market. In *6th International Conference on Control and Optimization with Industrial Applications (COIA)*. https://www.researchgate.net/publication/349297368_CRUDE_OIL_PRICE_FORECASTING_TECHNIQUES_IN_THE_WORLD_MARKET

Organization of the Petroleum Exporting Countries. (2022). *Annual report 2021*. https://www.opec.org/opec_web/static_files_project/media/downloads/publications/AR%202021.pdf

Radchenko, S. (2005). Oil price volatility and the asymmetric response of gasoline prices to oil price increases and decreases. *Energy economics*, *27*(5), 708-730. https://doi.org/10.1016/j.eneco.2005.06.001

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy economics*, *21*(5), 449-469. https://doi.org/10.1016/S0140-9883(99)00020-1

U.S. Energy Information Administration. (2023). *Short-term energy outlook September 2023*. https://www.eia.gov/outlooks/steo/pdf/steo_full.pdf

Wei, Y., Wang, Z., Li, D., & Chen, X. (2022). Can infectious disease pandemic impact the long-term volatility and correlation of gold and crude oil markets?. *Finance Research Letters*, *47*, 102648. https://doi.org/10.1016/j.frl.2021.102648

Yang, C. W., Hwang, M. J., & Huang, B. N. (2002). An analysis of factors affecting price volatility of the US oil market. *Energy economics*, *24*(2), 107-119. https://doi.org/10.1016/S0140-9883(01)00092-5

Yu, L., Zhang, X., & Wang, S. (2017). Assessing potentiality of support vector machine method in crude oil price forecasting. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(12), 7893-7904. https://doi.org/10.12973/ejmste/77926

Appendix A

Table A1

Statistical analysis summary

ARMA (p,q)	Statistically significant lags	SIGMASQ	Adjusted R^2	AIC	SBIC
(1,1)	1	0.0079	0.068	-1.980	-1,946
(1,2)	2	0.0079	0.070	-1.982	-1.948
(1,16)	1	0.0080	0.057	-1.968	-1.934
(1,18)	1	0.0079	0.063	-1.974	-1.941
(1,20)	2	0.0079	0.070	-1.982	-1.948
(3,1)	1	0.0079	0.070	-1.982	-1.948
(3,16)	1	0.0085	0.001	-1.911	-1.877
(3,18)	1	0.0084	0.006	-1.916	-1.822
(3,20)	2	0.0084	0.014	-1.923	-1.890
(4,1)	1	0.0078	0.076	-1.988	-1.954
(20,1)	2	0.0078	0.080	-1.993	-1.959
(20,2)	2	0.0084	0.010	-1.918	-1.886
(0,1)	1	0.0079	0.069	-1.983	-1.958

Vol. 22. № 3 (86). July–September 2023.
ISSN 2519-4070

Appendix B

Table B1

Comparison of Adjusted R2, AIC, and SIC for GARCH, EGARCH, and FIGARCH Models

	(1,2)	(1,20)
GARCH		
Adjusted R^2	0.067	0.070
AIC	-2.34	-2.35
SIC	-2.26	-2.29
EGARCH		
Adjusted R^2	0.065	0.069
AIC	**-2.36**	**-2.38**
SIC	**-2.29**	**-2.31**
FIGARCH		
Adjusted R^2	0.066	0.071
AIC	-2.32	-2.37
SIC	-2.25	-2.29

Appendix C

Table C1

The predictions of returns against their actual values

	RETURNWTI	RETURNWTIF
2022M03	0.183981	0.023918
2022M04	-0.061935	0.009047
2022M05	0.076341	-0.009172

Received: June 26, 2023.
Reviewed: July 28, 2023.
Accepted: August 10, 2023.