DIGITALES ARCHIV

Dritsaki, Chaido

Article

The performance of hybrid ARIMA-GARCH modeling and forecasting oil price

International Journal of Energy Economics and Policy

Provided in Cooperation with:
International Journal of Energy Economics and Policy (IJEEP)

The Performance of Hybrid ARIMA-GARCH Modeling and Forecasting Oil Price

Chaido Dritsaki*

Department of Accounting and Finance, School of Management and Economics, Western Macedonia University of Applied Sciences, Greece. *Email: dritsaki@teiwm.gr

ABSTRACT

Modeling and forecasting oil prices is an important issue for many researchers. One of the methods used in forecasting oil prices is Box-Jenkins methodology through ARIMA models. Although these models provide accurate forecasting over a short time period, they are not able to handle the volatility and nonlinearity presented on data series. For this reason, on this paper we examine a hybrid ARIMA-GARCH model in order to forecast the volatility in the return of oil prices. Moreover, on this paper, the Box-Cox transformation is used for data smoothing for the stabilization of variance and reduction of heteroscedasticity. Parameters' estimation in the hybrid ARIMA-GARCH model is employed by ML (Maximum Likelihood) method using the steps of Marquardt's Algorithm (1963) and Broyden-Fletcher-Goldfarb-Shanno algorithm for optimization. The results of the paper showed that the hybridation of ARIMA (33,0,14)-GARCH (1,2) model following normal distribution is the most suitable for forecasting the returns of oil prices. Finally, we use both the dynamic and static procedure for forecasting. The results showed that the static procedure provides with better forecasting than the dynamic.

Keywords: ARIMA, GARCH, Oil price forecasting, Hybrid ARIMA-GARCH, Box-Cox transformation
JEL Classifications: C33, O13, Q43

1. INTRODUCTION

Oil is considered to be one of the most important goods in the world. Its use is ubiquitous in everyday life. Due to its uniqueness, researchers need to develop a better understanding of its price dynamics so that industries that are supplied with or consume vast quantities of oil can take optimal decisions.

Oil's dynamic modeling is a difficult task to accomplish because its price cannot be predicted in various time periods and it depends on many factors. During the past decades, oil price shows volatility. In 1999, during the Asian crisis, Iraq's decision to increase oil production caused the decrease of oil price in the lowest level. In 2001, the dot-com bubble caused panic, reducing oil price until the beginning of 2002. Following this upheaval, global economy regained its momentum, resulting in an upward trend of oil price. The factors that led to the reduction of oil production were the hostile relationships between the U.S. with the production countries. Price oil reached its highest peak, when the housing bubble burst in the U.S. causing credit crisis. The reduction of oil

price that followed until its stabilization after the economic crisis of 2008, is a challenge for researchers investigating for a model who would forecast these situations.

Moreover, forecasting returns of oil prices, which is considered a basic economic variable, influences consumers' decisions, businesses and financial institutions, as well as governments. Timely and reliable predictions of oil prices provide important information on those in the financial markets.

This paper tries to develop a hybrid ARIMA-GARCH model in order to investigate and forecast the characteristics of volatility in oil price using daily data from 20 October 1997 until 31 May 2017 for a total of 4980 observations. The rest of the paper is as follows: Section 2 provides a brief literature review. Section 3 presents the analysis of methodology. Section 4 summarizes the data and the descriptive statistics. The empirical results are provided in Section 5 and Section 6 proposes the forecasting results. Finally, the last section offers the concluding remarks.

2. LITERATURE REVIEW

During last years, there is great interest from researchers for modeling volatility and forecasting oil price. Hansen and Lunde (2005) argued that information related to oil price is necessary for modeling producers' oil price as well as calculating risk measures. Also, others found out that variations in oil prices have consequences not only in those countries that import but also in those countries that export oil. Rothemberg and Woodford (1996) have studied oil price instability in relation to inflation. Hamilton and Herrera (2004) in their paper argue the relationship between oil price and exports. Yang et al. (2002) refer to the relationship between investment and oil price and Elder and Serletis (2009) examine the relationship between oil price and monetary policy.

Furthermore, there are a number of papers relating oil price with inflation, exchange rate, production and employment decrease and competitiveness loss. In all these papers, techniques using autoregressive models are employed (Mirmirani and Li, 2004), cointegration and error correction models (Mohammadi 2009), GARCH models (Sadorsky, 2006, and Agnolucci, 2009), as well as neural networks (Yu, et al. 2008).

Finally, it can be said that all papers reach the same conclusion that the returns of oil prices present a unit root, have excessive kurtosis, are negatively skewed and don't follow Gaussian distribution. Also, the volatility on the returns of oil prices is clustering and persistent, consistent with predictions of GARCH variety models.

3. THEORETICAL BACKGROUND

The development and designing of ARIMA models as forecasting tools of financial-economic variables is known as Box-Jenkins Methodology (1976). This methodology tries to find an ARIMA (p, d, q) model which satisfies the stochastic procedure where the sample derived from. Box-Jenkins methodology consists of four repetitive steps: Identification model, parameters' estimation, diagnostic tests and model's forecasting ability.

3.1. ARIMA Models
The autoregressive integrated moving average model of order p and q, ARIMA (p, d, q) is one of the time series forecasting methods for the nonstationary data series. The ARIMA (p, q, d) can be expressed as:

$$\phi_p(L)(1-L)^d(y_t-\mu)=\vartheta_q(L)e_t \tag{1}$$

or

$$\left(1-\sum_{i=1}^{p}\varphi_i L^i\right)\left(1-L\right)^d\left(y_t-\frac{1}{4}\right)=\left(1-\sum_{j=1}^{q}\vartheta_j L^j\right)e_t \tag{2}$$

Where,

$\phi_p\left(L\right)=1-\sum_{i=1}^{p}\phi_i L^i$ and $\vartheta_q\left(L\right)=1-\sum_{j=1}^{q}\vartheta_j L^j$ are polynomials in terms of L of degree p and q.

y_t is. the time series, and e_t is the random error at time period t, with μ is the mean of the model.

d is the order of the difference operator. $\varphi_1, \varphi_2,\ldots, \varphi_p$ and $\vartheta_1,\vartheta_2,\ldots, \vartheta_q$ are the parameters of autoregressive and moving average terms with order p and q respectively.

L is the difference operator defined as $\Delta y_t=y_t-y_{t-1}=(1-L)y_t$.

ARIMA models can be estimated following the Box-Jenkins approach. Given that the stationary procedure is essential for an ARIMA model, during the identification step data are transformed so that the time-series will become stationary. The stationary procedure is a necessary condition in building an ARIMA model.

3.2. Box-Cox Transformation Method
Box-Cox (1964) on their paper used a mathematical formula for the data transformation so that these can be more normally distributed and the variance equation corrects normality, linearity and reduces heteroscedasticity. The formula of the Box-Cox transformation is:

$$y_t^* = \begin{cases} \dfrac{y_t^\lambda -1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln\left(y_t\right) & \text{if } \lambda = 0 \end{cases} \tag{3}$$

Where, y_t
y_t are actual data in time t.
y_t^* are the transformed data in time t.

λ is the minimum value of mean square error of residuals.

The transformation of the above equation is valid only for positive values of time series $y_t>0$.

If the values of time series contain also negative values, then the transformation will get the following form:

$$y_t^*(\lambda) = \begin{cases} \dfrac{\left(y_t + \lambda_2\right)^{\lambda_1} - 1}{\lambda_1} & \text{if } \lambda_1 \neq 0 \\ \ln\left(y_t + \lambda_2\right) & \text{if } \lambda_1 = 0 \end{cases} \tag{4}$$

Where,
λ_1 is the transformation parameter, and
λ_2 is chosen such that $y_t>-\lambda_2$.

The main objective in the analysis of data transformation in Box-Cox (1964) technique is to calculate (estimate) λ parameter. For this reason, two approaches are necessary. The first approach is using the maximum likelihood method to estimate data because it facilitates the calculation of likelihood function. Also, maximum likelihood method is easy to obtain an approximate confidence interval for λ. The second approach uses Bayesian method to confirm if the model is fully specified.

3.3. GARCH Models
GARCH models are used mainly for modeling financial time series that present time-varying volatility clustering. The general

GARCH (r, s) model for the conditional heteroskedasticity according to Bollerslev (1986) has the following form:

$$y_t = \mu_t + z_t \tag{5}$$

Where,
μ_t is conditional mean of y_t.
z is the shock at time t.

$$z_t = \sigma_t e_t \tag{6}$$

Where,
$e_t \rightarrow iid\ N(0,1)$.

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i z_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2 \tag{7}$$

Where,

σ_t^2 is the conditional variance of y_t.

α_0 is a constant term.
r is the order of the ARCH terms.
s is the order of the GARCH terms.
α_i and β_i are the coefficients of the ARCH and GARCH parameters, respectively.

with constrains:

$$\begin{cases} \alpha_0 > 0 \\ \alpha_i \geq 0, \text{for } i = 1, 2, ..., r \\ \beta_i \geq 0, \text{for } i = 1, 2, ..., s \\ \sum_{i=1}^{r} \alpha_i + \sum_{i=1}^{s} \beta_i < 1 \end{cases}$$

3.4. Hybrid ARIMA-GARCH Model
In order to recommend a hybrid ARIMA-GARCH model, two stages should be applied. In the first stage, we use the best ARIMA model that fits on stationary and linear time series data while the residuals of the linear model will contain the non-linear part of the data. In the second stage, we use the GARCH model in order to contain non-linear residuals patterns. This hybrid model, which combines ARIMA and GARCH model containing non linear residuals patterns, is applied to analyze and forecast the returns of oil prices.

3.5. Estimation of Hybrid ARIMA-GARCH Model
The hybrid ARIMA-GARCH model is a non linear time series model which combines a linear ARIMA model with the conditional variance of a GARCH model. The estimation procedure of ARIMA and GARCH models are based on maximum likelihood method. Parameters' estimation in logarithmic likelihood function is done through nonlinear Marquardt's algorithm (Marquardt, 1963). The logarithmic likelihood function has the following equation:

$$\ln L[(y_t), \theta = \sum_{t=1}^{T} \left\{ \ln[D(z_t(\theta)), \upsilon] - \frac{1}{2} \ln[\sigma_t^2(\theta)] \right\} \tag{8}$$

Where θ is the vector of the parameters that have to be estimated for the conditional mean, conditional variance and density function, z_t denoting their density function, $D(z_t(\theta),\upsilon)$, is the log-likelihood function of $[y_t(\theta)]$, for a sample of T observation. The maximum likelihood estimator $\hat{\theta}$ for the true parameter vector is found by maximizing (8) (Dritsaki, 2017).

3.6. Diagnostic Checking of Hybrid ARIMA-GARCH Model
The diagnostic tests of hybrid ARIMA-GARCH models are based on residuals. Residuals' normality test is employed with Jarque and Bera (1980) test. Ljung and Box (1978) (Q-statistics) statistic for all time lags of autocorrelation is used for the serial correlation test. Also, for the conditional heteroscedasticity test we use the squared residuals of autocorrelation function.

3.7. Forecast Evaluation
On hybrid ARIMA-GARCH models we use both the static and dynamic forecast. The dynamic forecast, also known as n-step ahead forecast, uses the actual lagged value of Y variable in order to compute the first forecasted value. The static forecast (one-step ahead forecast) of Y_{t+1} based on an hybrid ARIMA-GARCH model is defined as:

$$\hat{Y}_t(1) = E\left(Y_{t+1} | Y_t, Y_{t-1}, ...\right) = \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t+1-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t+1-j} \tag{9}$$

Where the ε_s follow the stated GARCH model.

To evaluate the forecast efficiency, we use two statistical measures, mean squared error (MSE) and mean absolute error (MAE).

MSE it computes the squared difference between every forecasted value and every realised value of the quantity being estimated, and finds the mean of them afterwards.

MSE has the following formula:

$$MSE = \frac{1}{n} \sum_{i=1}^{n} \left(Y_i - \hat{Y}_i\right)^2 \tag{10}$$

Where,
Y_i is the vector of observed values of the variable being predicted.
$\hat{Y}_i$ is the vector of n predictions.

MAE it computes the mean of all the absolute, instead of squared, forecast errors. The formula is the following:

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left|Y_i - \hat{Y}_i\right| \tag{11}$$

4. DATA AND DESCRIPTIVE STATISTICS

The data used in our paper come from energy information administration. Data are daily covering the period 20 October 1997

until 31 May 2017 including 4980 observations. Using the value of λ = 0.3374 which was calculated from XLSTAT of Excel and Equation (3) from Box-Cox, we transformed oil price time-series.

In the following Figure 1, the closing values as well as transformed values in oil prices are presented.

From Figure 1 we can see that transformed data are less volatile from the initial ones. Then, we examine normality in oil prices before and after the transformation.

From Figure 2, we can see that the transformed data have a better adjustment in relation to normal distribution. In the following Table 1, the descriptive statistics of Brent index are presented before and after Box-Cox transformation.

The Table 1 show that standard deviation of transformed series has reduced from 33.78242 to 2.392691. Also, we can see that in the transformed data there is no normality even though it has been reduced.

The daily return of oil price index after the data transformation is calculated as follows:

$$r_t = \ln\left(\frac{y_t}{y_{t-1}}\right) \times 100 = \left[\ln y_t - \ln y_{t-1}\right] \times 100 \qquad (12)$$

Where,

y_t is the daily closing price of oil price index at day t.

y_{t-1} is the daily closing price of oil price index at previous day.

r_t are the daily returns of oil price index.

In the following Figure 3 we present the closing prices, returns and volatility of oil price index after the Box-Cox transformation. The volatility of oil prices is estimated from the daily squared returns (Sadorsky, 2006).

From Figure 3 we note that daily closing prices of oil follow a random walk whereas returns from oil prices seem to be stationary. The confirmation in stationary of the returns of oil price index is done with Dickey-Fuller (1979; 1981) and Phillips-Perron (1998) unit root tests.

The results of Table 2 confirm that returns of oil prices are stationary in their levels. Consequently, for an ARIMA (p, d, q) model the value for d = 0.

On Table 3, the descriptive statistics of the returns on Brent are presented.

The results on Table 3 show that the mean of daily return on oil price is quite small in relation to its standard deviation. Also, the return in oil prices appear small positive asymmetry and leptokurtosis with fat tails and Jarque and Bera (1980) statistic proves that the return of oil prices don't follow normal distribution.

Figure 1: Daily closing prices of oil before and after the transformation
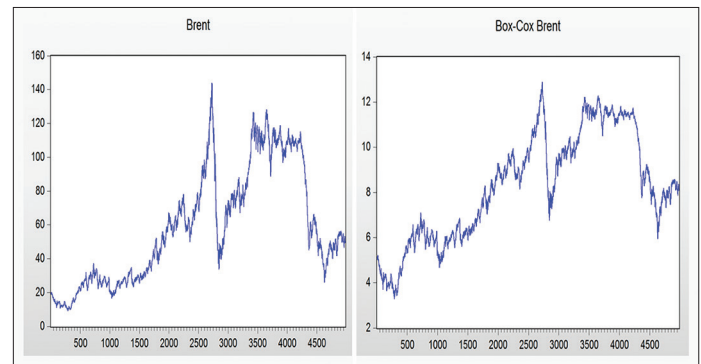
Figure 2: Normality test in daily prices of oil before and after transformation
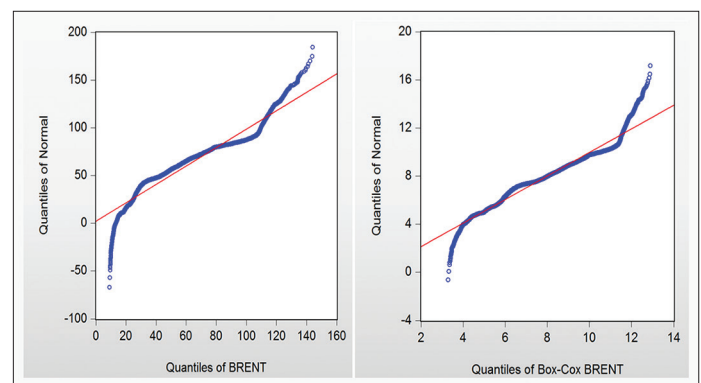
Table 1: Descriptive Statistics of Brent index before and after Box-Cox transformation

Brent		Box-Cox brent	
Mean	58.86256	Mean	8.281836
Median	53.15500	Median	8.362000
Maximum	143.9500	Maximum	12.88700
Minimum	9.100000	Minimum	3.280000
Standard deviation	33.78242	Standard deviation	2.392691
Skewness	0.436870	Skewness	−0.080189
Kurtosis	1.951764	Kurtosis	1.919715
Jarque-Bera	386.4104	Jarque-Bera	247.4930
Probability	0.000000	Probability	0.000000
Sum	293135.6	Sum	41243.54
Sum Sq. Dev.	5682294.0	Sum Sq. Dev.	28504.63
Observations	4980	Observations	4980

Table 2: Unit root tests of the returns of Brent

Brent	Augmented Dickey-Fuller		Phillips-Perron	
	C	C, T	C	C, T
	−69,3773*(0)	−69,3816*(0)	−69,3794*[9]	−69,3809*[8]

*,**,***show significant at 1%, 5% and 10% levels respectively. The numbers within parentheses followed by ADF statistics represent the lag length of the dependent variable used to obtain white noise residuals. The lag lengths for ADF equation were selected using Schwarz Information Criterion. MacKinnon (1996) critical value for rejection of hypothesis of unit root applied. The numbers within brackets followed by PP statistics represent the bandwidth selected based on Newey-West (1994) method using Bartlett Kernel. C=Constant, T=Trend

On Table 4, the autocorrelation diagram on the return of oil price is presented.

Figure 3: Closing prices, returns and volatility of brent index
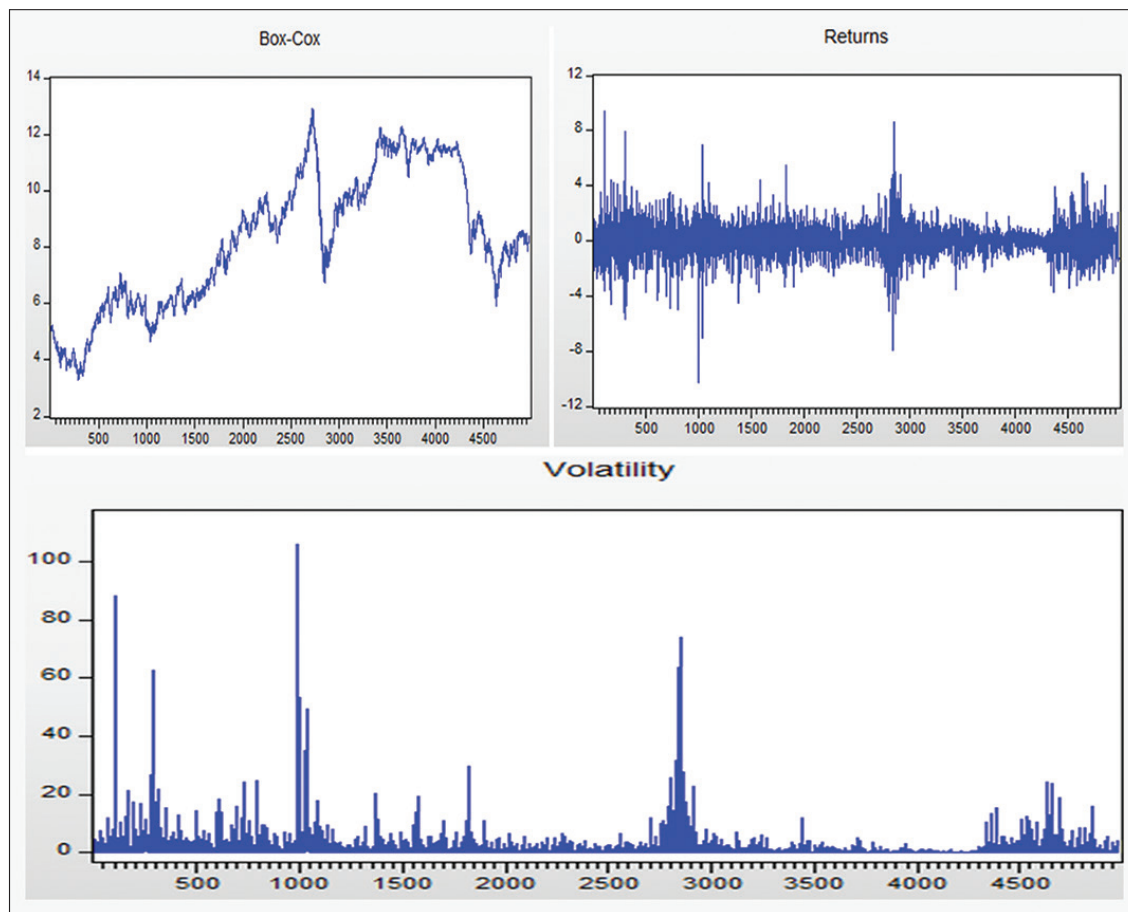
Table 3: Descriptive statistics of the returns of brent

Indice	Brent
Mean	0.0093
Median	0.0104
Maximum	9.3977
Minimum	−10.2945
Standard deviaton	1.1389
Skewness	0.0288
Kurtosis	9.0624
Jarque and Bera	7625.40
Probability	0.000000
Q (24)	54.619*
Observations	4979

*indicate statistical significance at 1%

The above results indicate that autocorrelation coefficients on lags 14,15,18,33 and 35 and partial autocorrelation on lags 14,15,18,28 and 33 on the return of oil price is larger than both standard errors $(\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{4980}} = \pm 0.02934)$. Furthermore, according to Ljung-Box statistic on Table 4, the existence of ARCH or GARCH cannot be rejected after the 13th lag.

5. EMPIRICAL RESULTS

Following, we define the form of ARIMA (p, d, q) model given the results of autocorrelation and partial autocorrelation diagram on Table 4. The p and q parameters of ARIMA model are determined from the coefficients of partial autocorrelation and autocorrelation respectively, comparing them with critical value $\pm \frac{2}{\sqrt{n}} = \pm \frac{2}{\sqrt{4980}} = \pm 0.02934$.

From the values of coefficients of partial autocorrelation and autocorrelation on Table 4 we see that the value of p will be p = 14 or p = 15 or p = 18 or p = 28 or p = 33 and for q will be q = 14, or q = 15 or q = 18, or q = 33 or q = 35. Using the above values, we choose the best ARIMA (p, 0, q) model from the smaller values of Schwarz criterion. Table 5 provides the values of p and q.

The results of Table 5 show that ARIMA (33,0,14) model is the most suitable for the returns of oil index. In the following Table 6 we get the estimations of this model.

Given the ARCH effects on the returns of oil price index, we proceed with the estimations of hybrid ARIMA-GARCH models to examine the volatilities that exist in the related returns of oil price. Moreover, from Figure 3 the returns in oil prices show cluster in volatility. To catch this cluster we should use ARIMA as well as GARCH models. Thus, in the levels this time-series on returns of oil prices we have to find out the appropriate hybrid ARIMA-GARCH model. Estimation parameters' is held with Maximum Likelihood method using the steps of Marquardt's algorithm (1963) and also Broyden-Fletcher-Goldfarb-Shanno algorithm optimization. Estimation parameters' as well as diagnostic tests

Table 4: Correlogram on the return of oil prices

Autocorrelation	Partial correlation	S. no	AC	PAC	Q-Stat	Prob
		1	0.017	0.017	1.3596	0.244
		2	0.006	0.006	1.5696	0.456
		3	0.005	0.005	1.6877	0.640
		4	−0.009	−0.009	2.0574	0.725
		5	0.000	0.000	2.0576	0.841
		6	−0.029	−0.029	6.3057	0.390
		7	0.024	0.025	9.1804	0.240
		8	0.008	0.007	9.5011	0.302
		9	0.013	0.013	10.340	0.324
		10	−0.005	−0.006	10.458	0.401
		11	−0.023	−0.023	13.065	0.289
		12	−0.011	−0.011	13.654	0.323
		13	0.018	0.020	15.226	0.293
		14	0.054	0.053	29.679	0.008
		15	0.036	0.034	36.041	0.002
		16	0.012	0.009	36.787	0.002
		17	−0.011	−0.014	37.393	0.003
		18	−0.035	−0.034	43.354	0.001
		19	−0.013	−0.009	44.149	0.001
		20	−0.021	−0.017	46.445	0.001
		21	−0.024	−0.024	49.405	0.000
		22	−0.013	−0.016	50.268	0.001
		23	0.008	0.005	50.574	0.001
		24	0.028	0.027	54.619	0.000
		25	0.014	0.017	55.558	0.000
		26	0.013	0.015	56.405	0.001
		27	−0.002	−0.003	56.417	0.001
		28	−0.027	−0.032	60.169	0.000
		29	−0.008	−0.012	60.483	0.001
		30	0.010	0.010	61.023	0.001
		31	−0.002	−0.001	61.035	0.001
		32	−0.013	−0.011	61.949	0.001
		33	−0.053	−0.052	76.224	0.000
		34	−0.004	0.000	76.292	0.000
		35	−0.031	−0.024	81.064	0.000
		36	0.003	0.010	81.100	0.000

Table 5: Comparison of ARIMA models within the range of exploration using Schwarz criteria

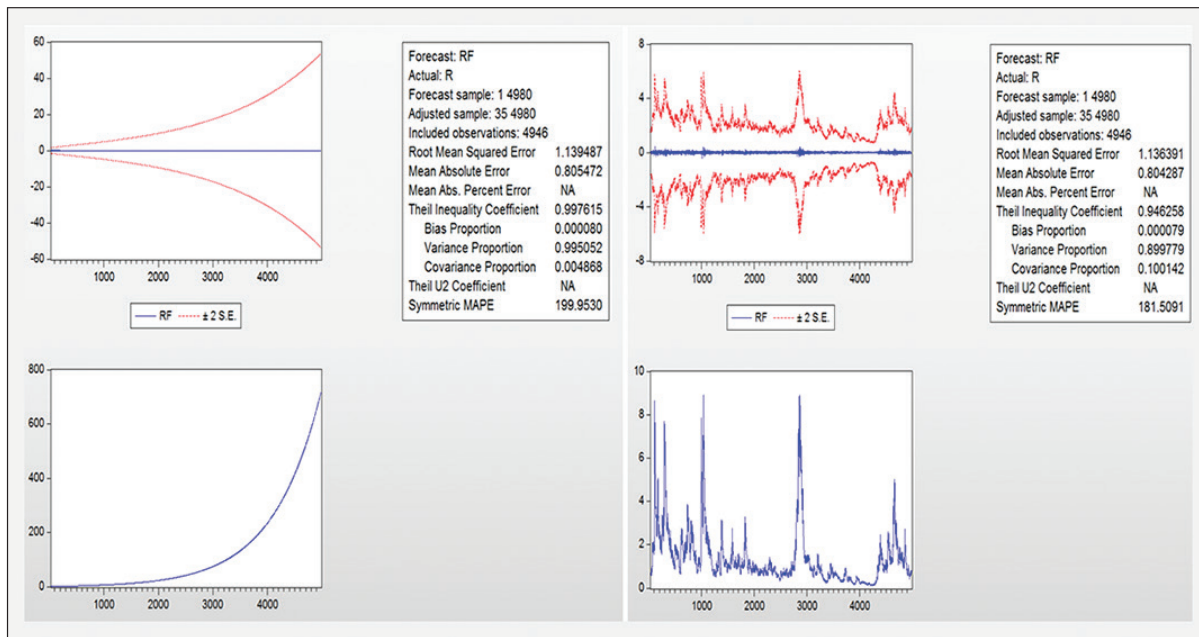| Returns of Brent	
ARIMA (14,0,14)	3.098784
ARIMA (14,0,15)	3.098888
ARIMA (14,0,18)	3.099006
ARIMA (14,0,33)	3.097087
ARIMA (14,0,35)	3.099275
ARIMA (15,0,14)	3.098691
ARIMA (15,0,15)	3.101659
ARIMA (15,0,18)	3.100680
ARIMA (15,0,33)	3.098897
ARIMA (15,0,35)	3.100866
ARIMA (18,0,14)	3.098830
ARIMA (18,0,15)	3.100695
ARIMA (18,0,18)	3.101842
ARIMA (18,0,33)	3.099034
ARIMA (18,0,35)	3.100882
ARIMA (28,0,14)	3.099227
ARIMA (28,0,15)	3.101020
ARIMA (28,0,18)	3.101108
ARIMA (28,0,33)	3.099381
ARIMA (28,0,35)	3.101431
ARIMA (33,0,14)	3.096988
ARIMA (33,0,15)	3.098987
ARIMA (33,0,18)	3.099104
ARIMA (33,0,33)	3.099902
ARIMA (33,0,35)	3.099255

Table 6: Estimations of the ARIMA (33,0,14) model of the returns of Brent

Variables	ARIMA (33,0,14)
AR (33)	−0.054488 (0.000)
MA (14)	0.058173 (0.000)
SIGMASQ	1.289110 (0.000)
Log likelihood	−7697.181
Jarque and Bera	7443.852 (0.000)
Q^2 (5)	414.48 (0.000)
X^2 (10)	329.4274 (0.000)
X^2 (20)	387.4811 (0.000)
X^2 (30)	407.5273 (0.000)

AR and MA denote the autoregressive and moving average terms respectively. SIGMASQ is the coefficient of variance error. $Q^2(5)$ is the Q-Statistic of correlogram of squared residuals at fifth lag. X^2 is the value of Chi-square of ARCH test and (10), (20), (30) are the corresponding lags. P values in parentheses denote probability.

of normality, autocorrelation and conditional heteroscedasticity are presented on Table 7.

From the Table 7 we can see that the hybrid ARIMA (33,0,14)-GARCH (1,1) with normal distribution is the most appropriate. All coefficients are statistical significant and there seems no problem in diagnostics tests (except normality). Persistence gets the value 0.99996 indicating high persistence in volatility in the oil price index.

Figure 4: Dynamic and static forecast of the ARIMA (33,0,14) - GARCH model of the returns of brent

Table 7: Estimates of the ARIMA (33,0,14) - GARCH models of the returns of Brent

Distribution	ARCH (1)			GARCH (1,1)		
	Normal	**t-student**	**GED**	**Normal**	**t-student**	**GED**
Mean equation						
AR (33)	−0.048140*	−0.031854*	−0.02396**	−0.03227**	−0.02503**	−0.021140
MA (14)	0.055167*	0.059645*	0.051390*	0.043054*	0.046523*	0.045317*
Variance equation						
α_0	1.070661*	1.051116*	1.015478*	0.002279*	0.001956*	0.002073**
α_1	0.176842*	0.232728*	0.204613*	0.055819*	0.041448*	0.047791*
β_1	---	---	---	0.944145*	0.958353*	0.951375*
T-Dist. DOF/GED parameter		4.144466*	1.101958*	---	6.977499*	1.389278*
Diagnostic tests						
Persistence	---	---	---	0.99996	0.999801	0.999166
Log L	−7558.006	−7231.040	−7239.278	−6982.652	−6879.77	−6889.303
Q² (20)	399.13*	353.11*	368.90*	21.371	28.79***	24.18
ARCH (10)	164.15*	148.70*	153.64*	11.391	20.69**	15.37
Jarque and Bera	8807.26*	9342.78*	9181.98*	736.55*	913.20*	816.71*

The persistence is calculated as $(\alpha_1+\beta_1)$ for the GARCH model. Log L is the value of the logarithmic likelihood. Q² (20) is the Q-Statistic of correlogram of squared residuals at twenty lag. ARCH (10) represents the F-statistic of ARCH test at 10th lag. *,**,***indicate statistical significance at 1%, 5% and 10% respectively

6. FORECASTING

For the forecasting of hybrid ARIMA (33,0,14)-GARCH (1,1) model we use both the dynamic and static procedure. The dynamic procedure computes forecasting for periods after the first sample period, using the former fitted values from the lags of dependent variable and ARMA terms. This procedure is called n-step ahead forecasts. The static procedure uses actual values and not forecasted values of the dependent variable. This procedure is called one step-ahead forecast. In the following Figure 4, we present the criteria for the evaluation of forecasting the returns of oil price using the dynamic and static forecast respectively.

The Figure 4 indicates that the static procedure gives better results rather than the dynamic because both mean squared error and MAE are lower in the static process.

7. DISCUSSION AND CONCLUSION

This paper aims to create a hybrid model combining ARIMA model with GARCH models of high volatility in order to analyze and forecast the return of oil price. According to various papers, the returns of oil prices have unit root, excessive kurtosis and negative skewness so they don't follow the Gaussian distribution. Instead, the transformation of Box-Cox was used to smooth the data. This resulted in the stabilization of variance and the decrease in heteroscedasticity. The empirical results of the paper showed that the hybrid ARIMA (33,0,14)-GARCH (1,1) provides the optimal results and improves estimation and forecasting in relation to previous methods. In conclusion, the combination of robust and flexible linear ARIMA models and the power of non linear GARCH models in handling volatility and the risk return of oil price, made hybrid models to be the most suitable for analysis and forecasting of time series.

REFERENCES

Agnolucci, P. (2009), Volatility in crude oil futures: A comparison of the predictive ability of GARCH and implied volatility models. Energy Economics, 31(2), 316-321.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Box, G.E.P., Cox, D.R. (1964), An analysis of transformations. Journal of the Royal Statistical Society, Series B, 26(2), 211-252.

Box, G.E.P., Jenkins, G.M. (1976), Time Series Analysis. Forecasting and Control. San Francisco: Holden-Day.

Dickey, D.A., Fuller, W.A. (1979), Distributions of the estimators for autoregressive time series with a unit root. Journal of American Statistical Association, 74(366), 427-431.

Dickey, DA., Fuller, W.A. (1981), Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica, 49(4), 1057-1072.

Dritsaki, C. (2017), An empirical evaluation in GARCH volatility modeling: Evidence from the Stockholm stock exchange. Journal of Mathematical Finance, 7, 366-390.

Elder, J., Serletis, A. (2009), Oil price uncertainty. Energy Economics, 31(6), 852-856.

Hamilton, J.D., Herrera, A.M. (2004), Oil shocks and aggregate macroeconomic behavior: The role of monetary policy: A comment. Journal of Money, Credit and Banking, 36(2), 265-286.

Hansen, P.R., Lunde, A.S. (2005), A forecast comparison of volatility models: Does anything beat a GARCH(1, 1)? Journal of Applied Econometrics, 20(7), 873-889.

Jarque, C., Bera, A. (1980), Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters, 6, 255-259.

Ljung, G.M., Box, G.E.P. (1978), On a measure of a lack of fit in time series models. Biometrika, 65(2), 297-303.

MacKinnon J.G. (1996), Numerical distribution functions for unit root and cointegration tests. Journal of Applied Econometrics, 11(6), 601-618.

Marquardt, D.W. (1963), An algorithm for least squares estimation of nonlinear parameters. Journal of the Society for Industrial and Applied Mathematics, 11, 431-441.

Mirmirani, S., Li, H.C. (2004), A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. Advances in Econometrics, 19, 203-223.

Mohammadi, H. (2009), Electricity prices and fuel costs: Long-run relations and shortrun dynamics. Energy Economics, 31(3), 503-509.

Newey, W.K., West, K.D. (1994), Automatic lag selection in covariance matrix estimation. Review of Economic Studies, 61(4), 631-654.

Phillips, P.C., Perron, P. (1998), Testing for a unit root in time series regression. Biometrika, 75(2), 335-346.

Rothemberg, J.J., Woodford, M. (1996), Imperfect competition and the effects of energy price increases on economic activity. Journal of Money, Credit and Banking, 28(4), 549-577.

Sadorsky, P. (2006), Modeling and forecasting petroleum futures volatility. Energy Economics, 28(4), 467-488.

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.

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