Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

Forecasting Indian crude oil Price Using Hybrid ARIMA-GARCH Model

Dhawani Nilesh Shah,
Research Scholar, Department of Statistics,
Gujarat University,
Ahmedabad, Gujarat, India

Manishkumar Thaker
Associate Professor and HoD, Statistics Department, M.G.Science Institute,
Ahmedabad, Gujarat, India

Abstract:

Crude oil is a vital energy commodity for many sectors, and India's demand for crude oil and petroleum products has grown steadily over the past decade due to socioeconomic and technical factors. Hence, forecasting oil prices is essential. A commonly used forecasting method is the Box–Jenkins method, which uses ARIMA models. It can provide accurate forecasts over short periods, but it struggles to handle the volatility and nonlinearity of oil prices. Thus, this paper proposes a hybrid ARIMA–GARCH model to forecast oil prices. The data used for forecasting are monthly Indian crude oil prices from April 2000 to October 2023. The results show that the ARIMA(4,2,0)–GARCH(0,3) hybrid model is the most suitable for forecasting oil prices. The hybrid model provides better forecasts and is expected to be valuable to all stakeholders in the industry, including researchers, policymakers, and businesses.

Key Terms: Oil price forecasting, ARIMA, GARCH, Hybrid ARIMA-GARCH

1. INTRODUCTION

Crude oil is one of the most important energy commodities and has a significant influence on the global economy [2]. For a growing economy such as India, the importance of oil is indisputable; India is a significant importer of oil [5]. To meet the needs of the expanding population, it is necessary to have a cost-effective supply of oil, which serves as a driving force for the development of the nation. The volatile nature of oil and its fluctuating global prices not only affects energy security but also economic indicators such as inflation rates and fiscal policies. To navigate this intricate landscape, a model for forecasting oil prices is essential. This forecasting model will exert significant influence not only on consumer decisions but also on business strategies, financial institutions, and even government policies[1]

Forecasting Indian crude oil prices is a complex study, and the energy sector is affected by numerous factors such as geopolitical events, the global economy, environmental issues, political factors, and supply–demand dynamics [3]. These variables contribute to the volatile nature of oil prices and make them difficult to forecast.

Addressing these challenges requires sophisticated methods; hence, the choice of a hybrid model that combines autoregressive integrated moving average (ARIMA) and generalised autoregressive conditional heteroskedasticity (GARCH) models is not arbitrary. The ARIMA model helps to capture time series patterns, and the GARCH model helps to capture volatility and heteroscedasticity.[4]

The main objective of this research paper is to build a hybrid ARIMA-GARCH model to forecast oil prices and study its basic characteristics. The data used are Indian crude oil prices published by the Petroleum Planning and Analysis Cell(https://ppac.gov.in) from April 2000 to October 2023, comprising 283 data points. This research paper is divided into five sections. Section 2 provides a literature review, Section 3 discusses the methodology, and Section 4 summarises the analysis and descriptive statistics. Section 5 presents the empirical summary, and section 6 proposes the forecasting results. Finally, the last section concludes the paper and discusses the conclusion and the way forward. The complete analysis was performed using R Studio.

2. LITERATURE REVIEW

Since the last two decades, there has been huge interest from researchers in modelling the volatility of the dynamic time series of crude oil prices. Miao, Ramchander, Wang, and Yang's (2017) paper titled "Influential Factors in Crude Oil Price Forecasting" studies the key factors that play a significant role in predicting crude oil prices. It is a valuable contribution to the field of energy as it helps to identify the drivers and determinants that impact crude oil price forecasts.[9]

The paper by Sultan 2020 titled "Empirical Investigation of Relationship between Oil Price and Inflation: The Case of India" discusses the relationship between oil prices and inflation. It states that fluctuation in oil prices impacts inflation rates and provides insights into economic significance.[10]
Sahu, Bandopadhyay, and Mondal (2015) "Crude Oil Price, Exchange Rate, and Emerging Stock Market: Evidence from India" discusses the relationship between crude oil prices, exchange rates, and the stock market. It provides insights into how these various external factors influence India's stock market.[8] Furthermore, numerous studies relate oil prices to

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

inflation, decrease in employment, loss of competitiveness, exchange rates, and production.

In addition, research papers that explore data-driven methodologies, such as time series for forecasting crude oil prices, have been published. Shambulingappa's (2020) titled "Crude Oil Price Forecasting using Machine Learning" discusses various machine learning techniques for predicting crude oil prices. This enables the integration of contemporary methods in forecasting commodities in the energy sector.[7]

Bhattacharya and Ahmed (2018) titled "Forecasting Crude Oil Price Volatility in India using a Hybrid ANN-GARCH Model" discusses an innovative approach to predicting crude oil prices using a hybrid model that combines an artificial neural network (ANN) and generalised autoregressive conditional heteroskedasticity (GARCH). This research has contributed to the Indian crude oil market by providing valuable insights.[6]

Meena, Singh, and Singh (2014) titled "Forecasting Mustard Seed and Oil Prices in India Using ARIMA Model" discusses the autoregressive integrated moving average (ARIMA) model to predict mustard seed and oil prices in India. It has helped us understand price forecasting in dynamic sectors such as agriculture and the edible oil sector.[11]

However, very few research papers have discussed the hybrid model for forecasting Indian crude oil prices. This research seeks to provide a comprehensive approach to forecasting methods and contribute to the Indian energy sector. This study aims to bridge this gap by constructing a hybrid ARIMA– GARCH model for Indian crude oil prices.

3. RESEARCHMETHODOLOGY

This section discusses the process of building ARIMA, GARCH, and hybrid ARIMA–GARCH models. Figure 1 briefly shows the above-mentioned steps.

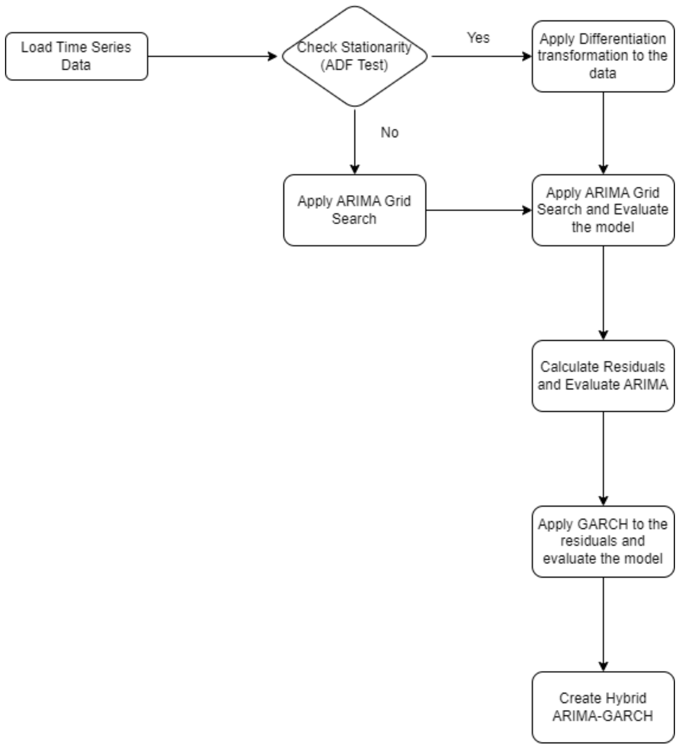

Figure 1:Brief Methodology of the Research Paper

3.1 ARIMA MODELS

The Box –Jenkins Methodology (1976) is an approach for developing and designing ARIMA models for forecasting.[12] This study aims to find an ARIMA model (p, d, q) that accurately represents the stochastic process in the sample data. The Box–Jenkins methodology has four steps: model identification, parameter estimation, diagnostic testing, and forecast evaluation.
The general form of the autoregressive integrated moving average represented as ARIMA (p, d, q) is expressed as follows:

$$\emptyset(L)(1-L)^d(y_t-\mu) = \vartheta_q(L)e_t \qquad (1)$$

Where

$\emptyset_p(L) = 1 - \sum_{i=1}^{p} \emptyset_i L^i$ and $\vartheta_q(L) = 1 - \sum_{j=1}^{q} \vartheta_j L^j$ are polynomials in terms of L of degrees p an q.
y_t is the time series z
e_t is the error term at time t
μ is the mean of the model
d is the order of difference
$\emptyset_p$ and ϑ_j are the parameters of autoregressive and moving average terms
L is a difference operator defined as $\Delta y_t = y_t - y_{t-1}$ = (1-L) y_t

The time series must be stationary in order to build an ARIMA model. In the situation where data are non-stationary, data are transformed so that the time series becomes stationary.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

3.2 GRID SEARCH METHOD

The Grid Search method is a comprehensive ML approach that rigorously explores a range of potential ARIMA model configurations that vary AR (p), differencing(d), and MA(q) parameters within a specified range. The method aims at minimising criteria such as the Aikaike Information Criterion (AIC) or Bayesian Information Criterion (BIC). This method ensures meticulous examination of all models, facilitating an informed selection of the ARIMA model that best captures all dynamics of crude oil price time series data.

3.3 HETEROSKEDASTIC MODELLING

The ARIMA model fails to consider the volatility of the variables. Generalized autoregressive conditional heteroskedasticity (GARCH) models are used to model financial time series with varying volatility Bollerslev's(1986) proposed GARCH(r,s) model[13], as given below:

$$y_t = \mu_t + z_t \tag{2}$$

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

$$z_t = \sigma_t \, e_t \tag{3}$$

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

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i \, z_{t-1}^2 + \sum_{i=1}^{s} \beta_i \, \sigma_{t-1}^2 \tag{4}$$

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.

3.4 HYBRID ARIMA– GARCH MODEL

The ARIMA model captures only the linear and stationary aspects of the data, whereas the residuals and the remaining nonlinear patterns are isolated. The hybrid ARIMA–GARCH model involves a two-stage process: initially, the ARIMA model is built, and then, in a subsequent stage, the GARCH model is built to analyse the residuals. This study addresses both the linear and nonlinear components of the time series.

4. DATA AND DESCRIPTIVE STATISTICS

The data under consideration comprises monthly crude oil prices of India published by the Petroleum Planning and Analysis Cell from April 2000 to October 2023, comprising 283 data points.
In Figure 1, crude oil prices are presented.

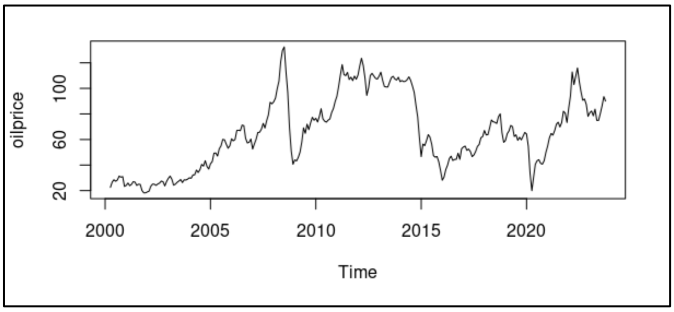

Figure 2: Monthly crude oil price

The descriptive statistics of the crude oil price are presented in Table 1. The results show that the mean oil price (65.04) is quite higher than the standard deviation (28.68) indicating the price volatility. The skewness of 0.25 says that the dataset is slightly positively skewed, implying longer tails, whereas the kurtosis of -0.97 indicates that the distribution has thinner tails than a normal distribution with fewer and fewer extreme outliers. Jarque and Bera (1980) statistics prove that oil prices do not follow a normal distribution.[14]

Crude Oil Price	
Mean	65.04
Standard deviation	28.68
Median	62.53
Minimum	18.24
Maximum	132.47
Skewness	0.25
Kurtosis	-0.97
Observations	283

Table 1: Descriptive Statistics of Crude Oil Price

To assess the stationarity of the time series, an augmented Dickey– Fuller (ADF) test was conducted. The results in Table 2 confirm that crude oil prices are non-stationary in their levels, and the probability value is greater than the significance level of 0.05 (5%). Thus, a differentiation transformation is performed to obtain the stationary time series data. This shows that later the model used is a model with the order d=1 or ARIMA(p,1,q).

Crude Oil Price	
Augmented Dickey-Fuller test	
Dickey-Fuller	-2.903
p-value	0.4959
Alternative Hypo-thesis	Stationary

Table 2: Test for Stationarity

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

5. EMPIRICAL RESULTS

Following, we define the form of the ARIMA (p,1, q) model given the results of the grid search method using AIC values. In consideration of the comprehensive scope, values of p and q are taken up to 20, Table 3 shows only the top 5 configurations that exhibited the most favorable model fits. This focused selection highlights the ARIMA configuration with the lowest AIC. The results in Table 3 show that the ARIMA (4,1,0) model is the most suitable for crude oil prices. This configuration implies that a first-order difference was necessary to achieve stationarity, and the model includes AR terms (p=4) without moving average terms (q=0). The AIC for the ARIMA (4,1,0) is 1803.977 and the RMSE is 5.87.

ARIMA Configuration	AIC
ARIMA (0,1,0)	1855.436
ARIMA (1,1,0)	1831.921
ARIMA (2,1,0)	1824.858
ARIMA (3,1,0)	1810.944
ARIMA (5,1,0)	1805.392
ARIMA (4,1,0)	1803.977

Table 3: Top 5 configurations of the Grid Search Result

Figure 3 shows the residual of the ARIMA (4,1,0) model. The autocorrelation for residuals indicates the volatility in the dataset and the need for the GARCH model. Given the ARCH effects on crude oil prices, we proceed with the estimations of hybrid ARIMA-GARCH models to examine the volatilities in the prices. Moreover, the residual indicates a cluster of volatility. To identify this cluster, we used the ARIMA and GARCH models. Thus, in the levels of this time series on crude oil prices, we find the appropriate hybrid ARIMA– GARCH model.

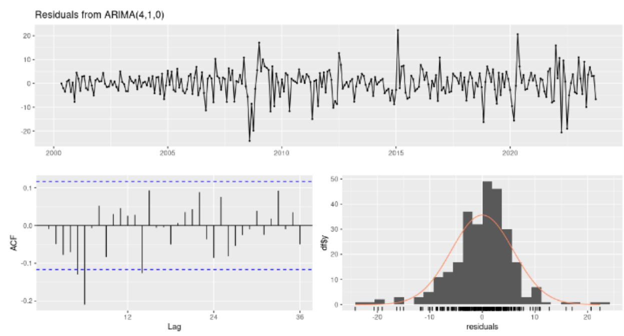

Figure 3:Residual plot of ARIMA(4,1,0)

From the ACF and PACF plots of the Squared Residuals and Grid method, the most suitable Hybrid model is ARIMA(4,1,0)-GARCH(0,3). Table 4 shows the coefficients of the Hybrid Model.

ARIMA (4,1,0)	
	Co-efficient
AR 1	-0.4320
AR2	-0.3069
AR3	-0.3042
AR4	-0.1771
GARCH (0,3)	
	Co-efficient
Mu	33.73
Omega	24.54
Beta 1	5.4e-07
Beta 2	0.00392
Beta 3	0.995

Table 4: Coefficient Table for Hybrid ARIMA(4,1,0)-GARCH(0,3*)*

6. FORECASTING

After obtaining results from the ARIMA (mean modelling) and GARCH (variance modelling), crude oil price is predicted for the next 1 month. From the above modelling results, a model that is good enough to represent Indian crude oil price data is the ARIMA (4,1,0)-GARCG (0,3) model. The forecast for November 2023 is 87.26, which is very close to the actual value of 87.09. This suggests that the model has less than 1% error.

7. DISCUSSION AND CONCLUSION

This study aims to create a hybrid model that combines the ARIMA model with the GARCH model of high volatility to analyse and forecast crude oil prices. The empirical results and forecasting values show that the hybrid ARIMA(4,1,0)-GARCH(0,3) method provides optimal results and improves estimation and forecasting in relation to previous methods. In conclusion, the combination of robust and flexible ARIMA models and the power of nonlinear GARCH models in handling the volatility of crude oil prices made the hybrid model the most suitable for analysis and forecasting time series.

8. REFERENCES

1. Sarmah, A., & Bal, D. P. (2021). Does crude oil price affect the inflation rate and economic growth in India? A new insight based on structural VAR framework. The Indian Economic Journal, 69(1), 123-139.
2. He, Y., Wang, S., & Lai, K. K. (2010). Global economic activity and crude oil prices: A cointegration analysis. Energy Economics, 32(4), 868-876.
3. Gkillas, K., Manickavasagam, J., & Visalakshmi, S. (2022). Effects of fundamentals, geopolitical risk and expectations factors on crude oil prices. Resources Policy, 78, 102887.
4. Mohammadi, H., & Su, L. (2010). International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. Energy Economics, 32(5), 1001-1008.

Li, S., & Wang, Q. (2019). India's dependence on foreign oil will exceed 90% around 2025-The forecasting results based on

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 11, November-2023

two hybridized NMGM-ARIMA and NMGM-BP models. Journal of Cleaner Production, 232, 137-153.

5. Bhattacharya, S., & Ahmed, A. (2018). Forecasting crude oil price volatility in India using a hybrid ANN-GARCH model.

7. Sahu, T. N., Bandopadhyay, K., & Mondal, D. (2015). Crude oil price, exchange rate and emerging stock market: Evidence from India. Jurnal Pengurusan, 42, 75-87.
8. Miao, H., Ramchander, S., Wang, T., & Yang, D. (2017). Influential factors in crude oil price forecasting. Energy Economics, 68, 77-88.
9. Sultan, Z. A. (2020). Empirical investigation of relationship between oil price and inflation: The case of India. International Journal of Energy Economics and Policy.
10. Meena, D. C., Singh, O. P., & Singh, R. A. K. E. S. H. (2014). Forecasting mustard seed and oil prices in India using ARIMA model. Annals of Agri-Bio Research, 19(1), 183-189.

International Journal of Business Forecasting and Marketing Intelligence, 4(4), 446-457.

6. Shambulingappa, H. S. (2020). Crude oil price forecasting using machine learning. International Journal of Advanced Scientific Innovation, 1(1), 1-11.
11. Box, G.E.P. and Jenkins, G.M. (1976) Time Series Analysis: Forecasting and Control. Revised Edition, Holden Day, San Francisco
12. Bollerslev, Tim, 1986. "Generalized autoregressive conditional heteroskedasticity," Journal of Econometrics, Elsevier, vol. 31(3), pages 307-327, April.
13. JARQUE, Carlos M. and BERA, Anil K. (1980) Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals. Economics Letters 6, 255-259.

