Journal of Basics and Applied Sciences Research (JOBASR)
ISSN (print): 3026-9091, ISSN (online): 1597-9962
Volume 4(1) January 2026
DOI: https://dx.doi.org/10.4314/jobasr.v4i1.11

Volatility Modelling of Crude oil Prices Using a Five-States Figarch-Hidden Markov Model Frame Work

Aliyu Mansur[1*], Ibrahim Lawal Kane[2], Abdulhameed Ado Osi[3], Huzaifa Abdurrahman[4], Hussaini Abubakar[5] & Usman Abubakar[6]
[1,2,3,4,5&6]Faculty of Computing and Mathematical Sciences, Department of Statistics, Aliko Dangote University of Science and Technology, Wudil
[*]Corresponding Author Email: **aliyumansur2016@mail.com**

ABSTRACT
Nigerian crude oil prices suffer from long memory, volatility, and varying volatility levels including extremely low, low, moderate, high, and extremely high. These characteristics directly increase market risk for the Nigerian economy. Studies model these levels into two levels low and high. However, the volatility levels can fall into other levels that include; extremely low, moderate, and extremely high. Therefore, this study aims to model these features using a newly developed hybrid time series model 5-States-FIGARCH-HMM. The data for this study was accessed documented record of central bank of Nigeria. The data were recorded monthly from 1990-2025 daily. The study employed ADF and KPSS tests to check for stationarity and non-unit roots was found. The study employed GPH test to test for presence of long memory in the time series and the long memory was found. Stationarity was achieved through fractional differencing. The study employed ACF and PACF plots to estimate the orders of AR and MA models respectively. The study found that 5-States-FIGARCH (1, 2)-HMM was the best model with least MAE, MSE, and RMSE when compared with FIGARCH and HMM models. The forecast results indicates rapid volatility stabilization and reduced tail risks, with a higher probability of low to moderate volatility regimes, implying more predictable pricing and reduced hedging requirements. Overall, the findings reveal strong regime persistence under calm market conditions, which may support macroeconomic stability toward early 2026.

Keywords:
HMM,
FIGARCH,
States,
Transition Probability,
Initial States Distribution,
Emission Parameters,
Volatility,
Regime Switching.

INTRODUCTION

Crude oil volatility refers to the fluctuations in crude oil prices over time. Understanding crude oil prices is crucial for investors, policymakers, and industry stakeholders due to its significant impact on the global economy (Kilian, 2022). Various factors contribute to crude oil volatility, including geographical events, supply and demand imbalances, and macroeconomic conditions (Ellwanger et al., 2023). Researchers have employed various models to capture crude oil volatility, such as GARCH-type models and stochastic volatility models (Hartwig and Mahringer, 2023). Crude oil volatility can be categorized into different levels, including; extremely low, low, moderate, high, and extremely high.

Low volatility characterized by small price fluctuations, low volatility is often observed during periods of economic stability and stable oil demand (Zhang et al., 2022). Moderate volatility represents moderate price fluctuations, often driven by changes in supply and demand fundamentals (Cross, and Nguyen, 2023). High volatility characterized by large price fluctuations, high volatility is often observed during periods of economic uncertainty, geopolitical tensions, and supply disruptions (Ellwanger et al., 2023). Extremely low volatility represents small prices fluctuations, often observed during periods of stable economic condition and low uncertainty (Barroso and Detzel, 2021). Extremely high volatility characterized by large price fluctuations, often driven by major shocks, such as wars, natural disasters, or global economic crises (Caldara et al., 2020). Understanding these different levels of volatility is essential for investors, policymakers, and industry stakeholders to make informed decisions and manage risk.

How to cite this article: Aliyu M., Ibrahim L. K., Abdulhameed A. O., Huzaifa A., Hussaini A. & Usman A. (2026). Volatility Modelling of Crude oil Prices Using a Five-States Figarch-Hidden Markov Model Frame Work. *Journal of Basics and Applied Sciences Research*, *4*(1), 97-112. https://dx.doi.org/10.4314/jobasr.v4i1.11

Recent studies have focused on the impact of COVID-19 on crude oil volatility, highlighting the increased uncertainty and price fluctuations during the pandemic (Sharif et al. 2020). Additionally, researchers have explored the relationship between crude oil volatility and other financial markets, such as stock markets and exchange rates (Yin et al., 2024).

Regime switching models are used to capture changes in the underlying data-generating process over time. These models assume that the data follows different regimes or states, each with its own set of parameters (Hamilton, 1989). Regime switching models have been widely applied in various fields, including finance and economics, to capture nonlinear relationships and structural breaks (Semmler and Toure, 2024). In the context of crude oil markets regime switching models have been used to capture changes in volatility and price dynamics (Mehrdoust et al., 2024). Recent studies have employed regime switching models to examine the impact of geographical events and economic policy uncertainty on crude oil prices (Yu et al., 2023).

Long memory refers to the persistence of shocks in the time series data over long periods. In the context of crude oil markets, long memory implies that shocks to oil prices can have lasting effects on future prices (Tiwari and Umar, 2021). Researchers have employed various models to capture long memory in crude oil prices, including fractional integration and long-memory GARCH models (Baillie et al., 1996).

Hybrid time series models combine different modelling approaches to capture complex patterns in data. These models have been increasingly used in various fields, including finance and capture nonlinear relationships. In the context of crude oil markets, hybrid models have been used to combine the strength of different modelling approaches, such ARIMA and GARCH models (Wang et al., 2018). Recent studies have employed hybrid models to examine the impact of various factors on crude oil prices, including economic policy uncertainty and geographical events (Balcilar et al., 2020). Hybrid models offer several advantages over standalone models, including improved forecasting accuracy, better capture of nonlinear relationships, and enhanced risk management capabilities (kumar et al., 2025). By combining different modelling approaches, hybrid models can capture complex patterns in data and provide more accurate forecasts (Burhan and Mohammed, 2024).

Many studies were carried out to model and analyse the impact of volatility on crude oil prices using the extension of HMM model. For example Deng et al. (2019) proposed an extension of HMM to Dynamic Time Wrapping and Hidden Markov Model (DTW-HMM) for forecasting and trading in crude oil market. The study applied DTW algorithm to match similar price sequences which have the same market state in historical time series, and then to calculate expected returns, while HMM approach was applied to classify time series into different states based on their development characteristic. The study used daily crude oil spot price from January 2, 1986, to December 30, 2017, as well as the daily price data of Brent crude oil spot price from May 20, 1987, to December 30, 2017. Experimental results showed that the proposed method yielded the best forecasting and trading performances in average. In the WTI market, the proposed method produced a hit ratio of about 62.74% and a yield of 34.3% profit per year, and a Sharpe ratio value of 2.274. Furthermore, experimental results of the proposed method were significantly superior to other benchmark methods, demonstrating that the proposed method is not only good at direction prediction and profit making, but also return/risk ratio. Moreover, Sengupta et al. (2023) developed a hybrid Hidden Markov Model-Long Short Term Memory (HMM-LSTM). The LSTM was employed to capture the complex patterns and dependencies. The models were employed to predict fluctuations in traffic flow, specifically the change in flow between successive time steps, instead of directly predicting the absolute flow values at a detector location. The performance evaluation of the proposed models was conducted on a dataset obtained from the California Department of Transportation's Performance Measurement System. Results indicated significant performance gained in using hybrid architecture compared to conventional methods such as Markov switching ARIMA and LSTM. Furthermore, Chenxing and Qiao (2025) introduced a novel Bayesian time series model that combined the nonparametric features of an Infinite Hidden Markov Model (IHMM) with the volatility persistence captured by the GARCH framework, to effectively model and forecast short-term interest rates. The model was applied to US 3-month Treasury bill rates. The GARCH-IHMM revealed both structural and persistent changes in volatility, thereby enhancing the accuracy of density forecasts compared to existing benchmark models. Out-of-sample evaluations demonstrated the superior performance of their model in density forecasts and in capturing volatility dynamics due to it is adaptively to different macroeconomic environments.

This study attempts to develop a hybrid time series model 5-states-Fractional Integrated Generalized Autoregressive Conditional Heteroscedasticity-Hidden Markov Model (5s-FIGARCH-HMM) that is capable to capture long memory, heteroscedasticity, regime switching and volatility levels. The model addresses the limitation of Generalized Autoregressive Conditional Heteroscedasticity-Infinite Hidden Markov Model (GARCH-IHMM), where the model cannot capture long memory and has no fixed number of states. The aim of the study is to develop a 5-State-FIGARCH-HMM that accurately captures the complex dynamics of crude oil prices in Nigeria, including long memory, heteroscedasticity, volatility, and regime switching. And the objectives are to: (i) develop a novel hybrid model that incorporates long memory, regime switching features, and volatility levels to capture the dynamics of crude oil prices in Nigeria (ii) apply the developed model to empirical data on crude oil prices in Nigeria to evaluate its performance and accuracy (iii) compare the performance of the develop model with existing models (iv) forecast the volatility of crude oil prices using the proposed model and evaluate its performance using metrics.

MATERIALS AND METHODS

Model Specification: Five-State FIGARCH-HMM

This study employs a Five-State Fractional Integrated Generalized Autoregressive Conditional Heteroscedasticity-Hidden Markov Model (Five-State FIGARCH-HMM) to capture long memory, volatility clustering, and regime switching in crude oil price volatility.

The joint probability distribution of the observed series and hidden state is defined as:

$$P(Z_1, Z_2, \dots, Z_T \;\; X_1, X_2, \dots, X_T) = P(X_1).\prod[P(X_t|X_{t-1}).P(Z_t|X_t)] \quad (1)$$

Where,

Z_t : is the residual of FIGARCH model at time t.

$P(X_1)$: is the initial state distribution (π)

$P(X_t|X_{t-1})$: is the transition probability (A)

$P(Z_t|X_t)$: is the emission probability (B)

X_t : is the state at time t.

$X_t = \{1, 2, 3, 4, 5\}$.

5-State-FIGARCH-HMM Parameters Estimation

The likelihood

$$\mathrm{P}(Z1,\dots,Z_T,\; X_{1,\dots,}X_T) = P(X_1).\prod_{t=2}^{T} P(X_t/X_{t-1}) \left(\prod_{t=1}^{T} P(Z_T/X_T)\right) \quad (2)$$

This is the chance that both the whole hidden-state sequence $(X_{1,\dots,}X_T)$ and the whole observation sequence $(Z1,\dots,Z_T)$ occur together.

Where, $P(X_1)$ is the probability that the chain begins in state X_1, $\prod_{t=2}^{T} P(X_t/X_{t-1})$ is the probability of moving from the previous state to the next one for each step from 2 up to T and $(\prod_{t=1}^{T} P(Z_T/X_T)$ is the probability of observing (Z_T) given the hidden state at that time for each time point.

Transition Probabilities Matrix (A)

Transition probability matrix represents the probability of moving from one state to another. This is a matrix of probabilities that govern the transitions between all the states. It is defined as:

$$A = \{a_{ij}\} \quad (3)$$

Where,

$$a_{ij} = P(X_t = j|X_{t-1} = i) \quad (4)$$

Properties

i. $a_{ij} \geq 0 \forall i, j$

ii. $\sum_i a_{ij} = 1 \; \forall i$

Emission Probability Matrix (B)

Emission probability matrix represents the probability of observing a particular value given the current state. This is a matrix probability that govern the emission of observations given the state. it is defined below as:

$$B = \{b_{jk}\} \quad (5)$$

Where,

$$b_{jk} = P(O_t = k|X_t = j) \quad (6)$$

Properties

i. $b_{ij} \geq 0 \forall i, k$

ii. $\sum_k b_{ik} = 1 \; \forall j$

FIGARCH Model

The FIGARCH model introduced by Baillie et al (1996) is specified as:

i. Fractional Integration.
ii. Long memory effects.
iii. Conditional variance.
iv. Flexibility.
v. Heteroscedasticity.

The model is defined as:

$$\sigma_t^2 = \omega + \beta(L)\sigma_{t-1}^2 + [1 - \beta(L) - \emptyset(L)(1 - L)^d]\varepsilon_t^2 \quad (7)$$

Where,

σ_t^2 : is the conditional variance at time t.

ω : is the constant variance.

$\beta(L)$: is the lag operator polynomial for GARCH term.

$\emptyset(L)$: is the lag operator polynomial for ARCH terms.

d : is fractional differencing parameter.

ε_t^2 : is the squared residual.

L : is the lag operator.

σ_{t-1}^2 : is the past return of the conditional variance.

The standardized residuals of FIGARCH model are defined as:

$$\varepsilon_t = \sqrt{\frac{\sigma_t^2 - \omega - \beta(L)\sigma_{t-1}^2}{1 - \beta(L) - \emptyset(L)(1 - L)^d}} \quad (8)$$

Let, $\varepsilon_t = Z_t$ (9)

Hidden Markov Model

Hidden Markov Model was developed by Baum (1960s). The model is used to model the behaviour of a system that is not directly observable.

The model is defined as:

$$P(Y_1, Y_2, \dots, Y_T \;\; X_1, X_2, \dots, X_T) = P(X_1).\prod[P(X_t|X_{t-1}).P(Y_t|X_t)] \quad (10)$$

The equation (4) represents the joint probability distribution of the observations (Y) and the hidden states (X).
Where,
$P(X_1)$: is the initial state distribution (π)
$P(X_t|X_{t-1})$: is the transition probability (A)
$P(Y_t|X_t)$: is the emission probability (B)
By substituting equation (3) into (4) we get equation (5) which the 5s-FIGARCH-HMM

5-State-FIGARCH-HMM Parameters

Initial State Distribution (π)

Initial state distribution represents the probability of starting in each state at time t = 1. This is a probability distribution over the set of possible states. It is defined as:

$$\pi = \{\pi_i\} \tag{11}$$

Where,

$$\pi_i = P(X_1 = i) \tag{12}$$

Properties

i. $\pi_i \geq 0 \forall i$
ii. $\sum_i \pi_i = 1$

Transition Probabilities Matrix (A)

Transition probability matrix represents the probability of moving from one state to another. This is a matrix of probabilities that govern the transitions between all the states. It is defined as:

$$A = \{a_{ij}\} \tag{13}$$

Where,

$$a_{ij} = P(X_t = j|X_{t-1} = i) \tag{14}$$

Properties

iii. $a_{ij} \geq 0 \forall i, j$
iv. $\sum_i a_{ij} = 1 \ \forall i$

Emission Probability Matrix (B)

Emission probability matrix represents the probability of observing a particular value given the current state. This is a matrix probability that govern the emission of observations given the state. It is defined below as:

$$B = \{b_{jk}\} \tag{15}$$

Where,

$$b_{jk} = P(O_t = k|X_t = j) \tag{16}$$

Properties

iii. $b_{ij} \geq 0 \forall i, k$
iv. $\sum_k b_{ik} = 1 \ \forall j$

Method of Parameter Estimation of 5s-FIGARCH-HMM

This study employed Expectation-Maximization (EM) algorithm to estimate the parameters of Sthe developed model. The steps undertaken to estimate the parameters are as follows:

i. Initialize parameters (π, A, B).
ii. E-Step: calculate expected counts of state transitions and observations.
iii. M-Step: update parameters using expected counts.
iv. Repeat E-step and M-step until converge.

Diagnostic Tests

Diagnostic test is the process of evaluating the goodness of model. The study used Quantile to Quantile plot to evaluate the goodness of the developed model.

Quantile-Quantile Plot

This study used Quantile-Quantile plot as the graphical procedure to diagnose the normality of the residuals of the developed model.
The decision is based on the following hypotheses:
H_0: The residuals of the model are normally distributed.
H_a: The residuals of the model are not normally distributed.
Reject null hypothesis if the residuals not follow or approximately follow line of best fit otherwise accept.

Evaluation Metric Tests

The study used the following evaluation metric test to select the best model among the models and to assess the forecast power of the selected best model.

i- $$\text{MAE}=\frac{1}{n}\sum_{t=1}^{n}\left|\sigma_t^2-\hat{\sigma}_t^2\right| \tag{17}$$

ii- $$\text{MSE}=\frac{1}{n}\sum_{t=1}^{n}\left(\sigma_t^2-\hat{\sigma}_t^2\right)^2 \tag{18}$$

iii- $$\text{RMSE}=\sqrt{\frac{1}{n}\sum_{t=1}^{n}\left(\sigma_t^2-\hat{\sigma}_t^2\right)^2} \tag{19}$$

Long Memory Checking

Long memory is a phenomena when a time series exhibits decay slowly rather than exponential decay. This study employed GPH test introduced by Gewek and Porter Hudak (1983) to estimate and check for the long memory. The test is defined as:

$$\ln\left[I\left(w_j\right)\right]=\beta_0+\beta_1\ln\left[4\sin\left(\frac{w_j}{2}\right)\right]+\varepsilon_j \tag{20}$$

Where, $w_j=\frac{2\pi_j}{T}$, $j=1,2,\ldots,n$, w_j refers to Fourier frequency Transformation $\left(n=\sqrt{T}\right)$ ε_j represents residual of the model, $I\left(w_j\right)$ is a simple periodogram which is defined as:

$$I\left(w_j\right)=\frac{1}{2\pi T}\left|\sum_{i=1}^{T}\varepsilon_t\lambda^{-w_jt}\right| \quad (21)$$

The test is based on the following decision criterion:
d = 0: indicates no long memory (short memory).
0 < d < 0.5: indicates anti-persistent behavior.
d = 0.5: indicates random walk.
0.5 < d < 1: indicates persistent behavior (long memory)
d = 1: indicates a non-stationary process.

Heteroscedasticity Checking

Heteroscedasticity refers to the presence of non-constant variance in time series data. The study employed scatter plot to check for the presence of heteroscedasticity in the residuals of FIGARCH model.

Scatter Plot

This research work utilised scatter plot to check for the presence of heteroscedasticity in the residuals of FIGARCH model (1, 4), the residuals were plotted against the time. The residuals are on the vertical line and the times are on the horizontal line.

Stationarity Checking

Stationarity is the phenomenon when the mean and variance of time series data are constant. This study employed Augmented Dickey-Fuller (ADF) and Kwiat-Kowski Smidth-Shin (KPSS) test to check for stationarity in the currency in circulation time series data.

Augmented Dickey-Fuller Test

This study checked for unit roots in the time series with the test which was developed by Said and Dickey-Fuller (1984). The test is defined below as:

$$\tau=\frac{(\beta_1-d)}{\sqrt{\frac{\beta_2^2+\sigma^2}{(1-\beta^2)^2}}} \quad (22)$$

Where, β_1 is the trend term, d is the differencing parameter, β_2 is the coefficient on the lagged first difference term, and σ^2 is the variance of the time series.
The test involves the following hypotheses:
H_0: the time series has unit roots.

H_a: the time series non-unit roots.

Decision Criteria

Null hypothesis is rejected if P-value is less than the alpha value.

KPSS Test

Time Series Visualization

This research work checked for the presence of unit roots with the test which was proposed by Kwiat-Kowski Smidth-Shin (1992). The test serves as the second approach to check for the unit roots of the time series data. The test is defined below as:

$$KPSST=\frac{\sum_{t=1}^{T}(y_t-\hat{\mu}-\hat{\delta}_t)^2}{\hat{\sigma}^2\sum_{t=1}^{T}\left(1-\frac{t}{T}\right)^2} \quad (23)$$

Where, y_t is the time series, $\hat{\mu}$ is the mean of the time series, $\hat{\delta}_t$ is the estimated trend coefficient, $\hat{\sigma}^2$ is estimated variance of the time series, and T is the sample size.

Model Order Selection

This study used Autocorrelation and Partial Autocorrelation function plot to identify the order of the model developed model.

Autocorrelation Plot

The study employed autocorrelation plot to identify the order of Moving Average (MA) model to use in the developed model. The autocorrelation function is given below as:

$$\rho_k=\frac{\theta_k}{\theta_0} \quad (24)$$

Where, ρ_k is the autocorrelation at lag k, k is the chosen lag, θ_k is the covariance at lag k and θ_0 is the variance.

Partial Autocorrelation Plot

The study used Partial Autocorrelation Function (PACF) plot to identify the order of Autoregressive (AR) model to use in the proposed model. The function is defined as:

$$\varphi(k)=\frac{[\rho(k)-\Sigma[\varphi(j).\rho(k-j)]]}{[1-\Sigma[\varphi(j).\rho(j)]]} \quad (25)$$

Where, $\varphi(k)$ is the partial autocorrelation at coefficient at lag k, $\rho(k)$ is the autocorrelation coefficient at lag k, $\varphi(j)$ is the partial autocorrelation at lag j, k is the number of lag, and j is the intermediate lag (j = 1 to k – 1).

Data for Application

The study used monthly Nigerian crude oil prices data covering 1990-2025, obtained from the central bank of Nigeria (CBN).

RESULTS AND DISCUSSION

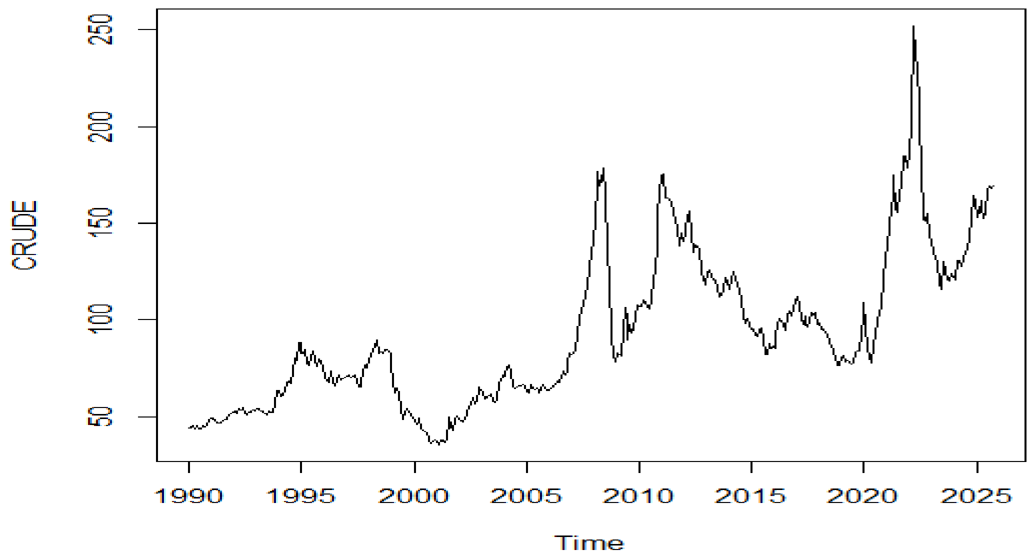


Figure 1 Time Series Plot of Crude oil

From the results obtained in Figure 1 it is observed that the time series behaviour exhibits an upward trend with evidence of structural breaks, indicating potential non-stationarity.

Stationarity Checking

Table 1 ADF Test Results

Dickey-Fuller = -3.5091	Lag order = 7	P-value = 0.04171

From the results obtained in Table 1 it is observed that the probability value of the ADF test is 0.04171, which is less the level of significance 0.05. Thus, we reject to the null hypothesis and conclude that the time series is stationary.

Table 2 KPSS Test Results

KPSS Level = 4.0833	Truncation lag parameter = 5	P < 0.01

From the results obtained in Table 2 it is observed that the probability value of KPSS test is less than 0.01, which is less than 0.05, thus, we fail to reject the null hypothesis and conclude that the time series is not stationary.

The stationarity checking suggests two distinct results, the ADF test suggests stationarity while the KPSS test suggests non-stationarity. Therefore, we are to conclude that the time series is not stationary.

Long Memory Checking

Table 3 GPH Test of DPI

Estimated (d)	sd.as	sd.reg
0.9060606	0.1812318	0.2040656

From the results obtained in Table 3 it is observed that the estimated long memory parameter d = 0.9060606, which suggest significant long memory in the time series data. This means the series has persistent, slowly fading shocks.

Time Series Visualization of the Differenced Data

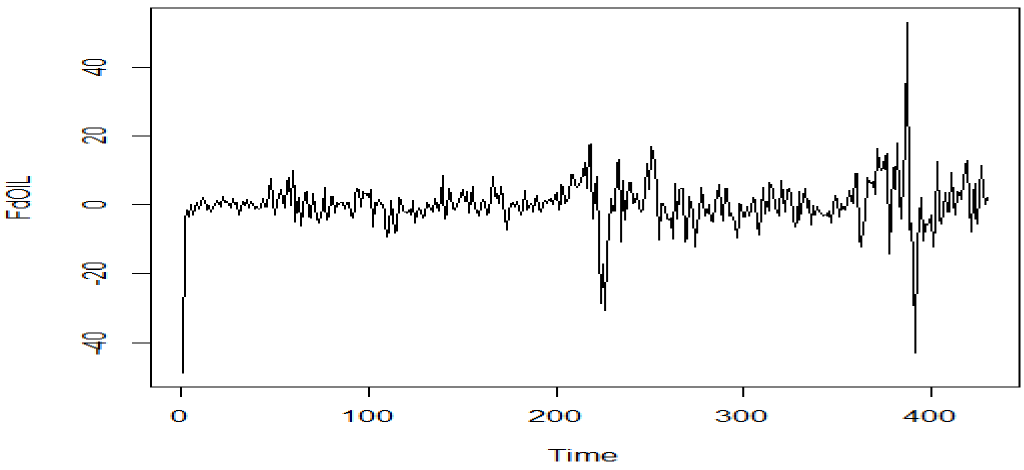


Figure 2 Time Series Plot of the Difference Time Series

From the results obtained in Figure 2 it is observed that the differenced series fluctuates around a constant mean with no visible trend, suggesting, stationarity after fractional differencing.

Stationarity Checking of the Differenced Time Series

Table 4 ADF Test Results of FdOIL Time Series

Dickey-Fuller = -6.7409	Lag order = 7	P-value = 0.01

From the results obtained in Table 4 it is observed that the probability value of ADF test is 0.01, which is less than 0.05, thus, the we are to reject the null hypothesis and conclude that the time series is stationary. However, this test is not enough to conclude that the time series is stationary.

Table 5 KPSS Test Results of FdOIL Time Series

KPSS Level = 0.11437	Truncation lag parameter = 5	P-value > 0.1

From the results obtained in Table 5 it is observed that the probability value of KPSS test is greater than 0.1, which is greater than 0.05, thus, we are to reject the null hypothesis and conclude that the time series is stationary.

Model Order Estimation

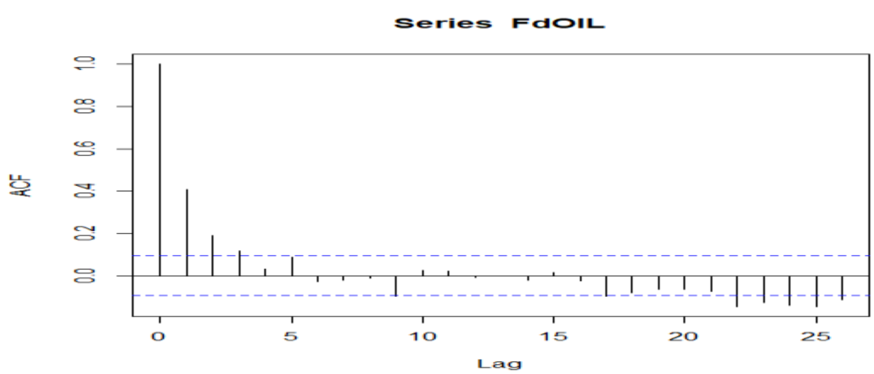
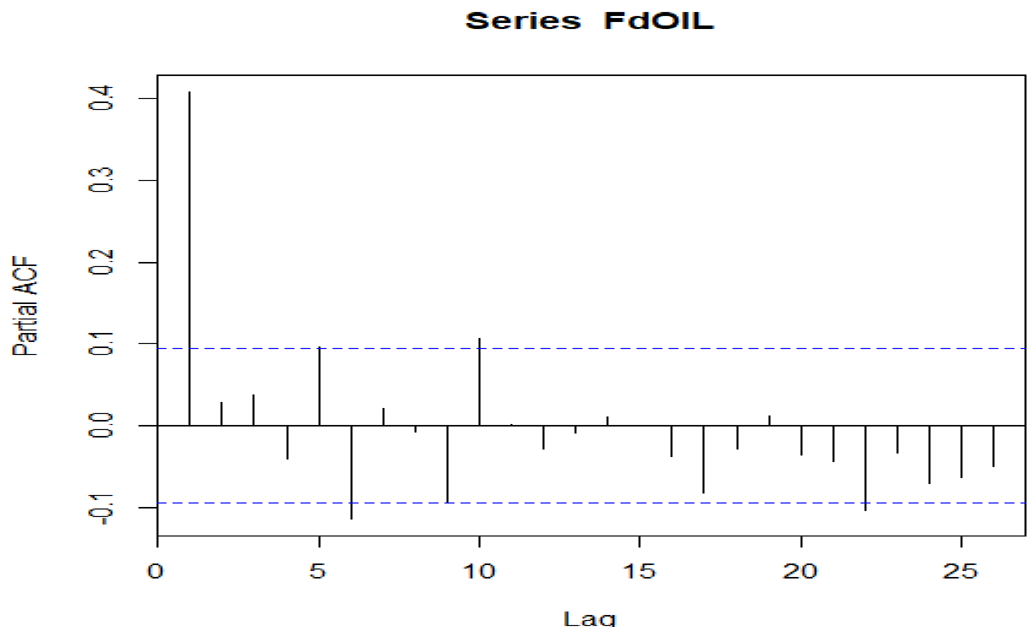


Figure 4.3 Autocorrelation Function Plot

From the above results obtained in Figure 3 it is observed that the autocorrelation plot exhibits significant spike from lag 1 up to lag 5. Thus, MA (1), MA (2), MA (3), MA (4), and MA (5) are significant.



4 Partial Autocorrelation Function Plot

From the results obtained in Figure 4.4 it is observed that the autocorrelation plot exhibits significant spike at lag 1. Thus, AR (1) is significant.

FIGARCH Model Estimation

Table 6 Information Criterion of FIGARCH Models

Model	MSE	RMSE
FIGARCH (1, 1)	49.21526	7.015359
FIGARCH (1, 2)	49.16552	7.011813
FIGARCH (1, 4)	49.12025	7.008584
FIGARCH (1, 5)	49.89544	7.063671

From the results obtained in Table 6 it is observed that the FIGARCH (1, 4) model outperforms the other models with least MSE and RMSE.

Heteroscedasticity Checking

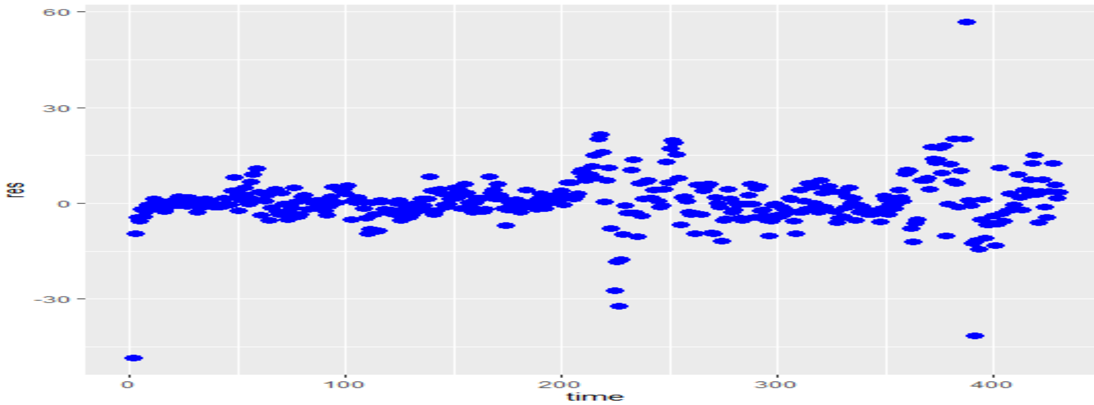


Figure 5 Volatility Plot

From the results obtained in Figure 5 it is observed that the residuals plot exhibits a fan shape pattern, confirming the presence of conditional heteroscedasticity, thereby justifying the use of a FIGARCH framework.

Assumptions of FIGARCH-5States-HMM

Table 7 Results Tests of the Assumptions

Presence long memory	Achieved	GPH Test d = 0.9
Stationarity	Achieved	ADF P.value < 0.05 KPSS P.value > 0.05
Markov property	Achieved	ACF shows decay slowly
Emission	Achieved	LRT P.value = 8.881784e-16
Transition probabilities	Achieved	Chi-square P.value = 2.2e-16
Initial state distribution	Achieved	Chi-square P.value = 0.406
Finite number of states	Achieved	Five (5)
Time invariant	Achieved	0.4762

From the results obtained in Table 7 it is observed that all the assumptions of FIGARCH-5States-HMM were achieved. Thus, the model is indeed a predictive model.

Table 8 Likelihood Ratio Test of Emission

LR	Df	Probability value
1.318261e+02	2.700000e+01	8.881784e-16

From the results obtained in Table 8 it is observed that the probability value of LRT is 8.881784e-16, which less than 0.05. This implies that s single emission distribution cannot adequately describe the volatility; instead, the five states exhibit statistically distinct emission parameters, confirming the presence of clearly separate volatility levels (extremely low to extremely high).

Table 9 Chi-square Test for Transition Probabilities

X-squared	Df	Probability value
1106.9	16.000	p-value < 2.2e-16

From the results obtained in Table 9 it is observed that the probability of the test is ($P < 0.05$), indicating a strong dependence between current and past states and supporting the Markov transition capture by the FIGARCH-5-State HMM.

Table 10 Chi-square Test for Initial Probability Distribution

X-squared	Df	Probability value
4.000	4.000	0.406

From the results obtained in Table 10 it is observed that the probability value is 0.406, this indicates that there is no evidence against the null hypothesis, meaning the initial probabilities fit the data well and appear stationary or consistent with historical patterns.

Table 11 OLS-based CUSUM test for Time Invariant

S0	p-value
0.75695	0.6155

From the results obtained in Table 11 it is observed that the test for time invariant yields p = 0.6155, failing to reject the null hypothesis of parameter stability. This confirms that the FIGARCH-5States-HMM parameters remain consistent over the sample period, supporting the model's applicability for analysing persistent volatility regimes in oil markets.

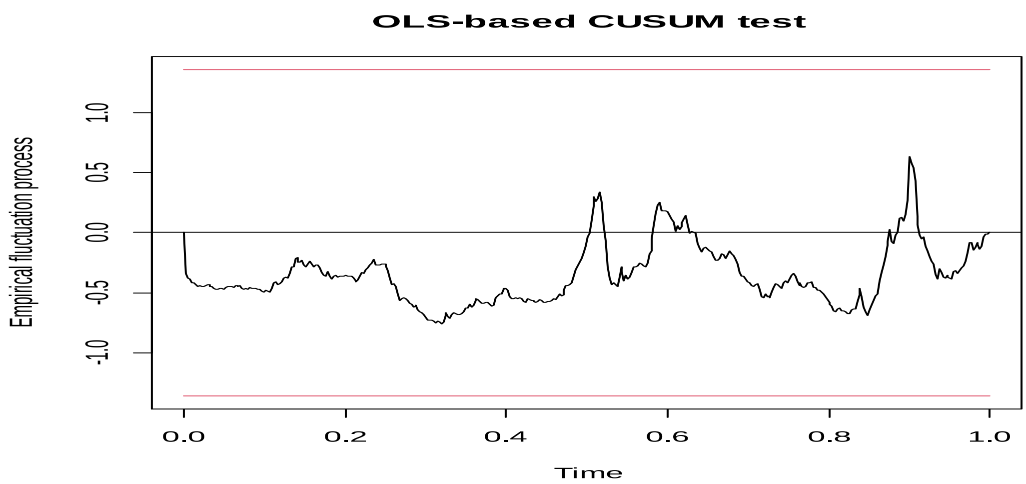


Figure 6 Parameters Stability Plot

From the above results obtained in Figure 6 it is observed that the plot's overall trend suggests parameters were stable early on but destabilized later, highlighting evolving regime dynamics. This is a standard output for econometric stability checks, emphasizing the need for break-aware modelling in volatile series like oil prices.

Hidden Markov Model

Table 12 Initial State distribution

Pr(extremely low)	Pr(low)	Pr(moderate)	Pr(high)	P(extremely high)
0	1	0	0	0

From the results obtained in Table 12 it is observed that since the value of probability 2 is 1, this indicates that the market is definitely at low volatility state at time zero. This means we start by believing the crude oil market is calm (low volatility) before any data are observed. The other states have zero chance initially. Hence, the market is in stable, low volatility phase before we observe any price changes.

Table 13 Transition Matrix

States	To S1	To S2	To S3	To S4	To S5
From S1	0.925	0.000	0.038	0.000	0.037
From S2	0.000	0.977	0.000	0.023	0.000
From S3	0.074	0.000	0.895	0.031	0.000
From S4	0.000	0.012	0.037	0.951	0.000
From S5	0.036	0.000	0.000	0.000	0.964

From the results obtained in Table 13 it is observed that from state 1 (extremely low volatility) mostly stays in state 1 (92.5%), but sometimes jumps to moderate and extremely high volatility (state 3 and 5, 3.8% and 3.7% respectively). The market tends to stay calm but can suddenly spike to moderate and extremely high volatility. Moreover, from state 2 (low volatility) it is observed that the volatility is very stable, staying 97.7% of the time, but small chance (2.3%) to jump suddenly to high volatility (state 4), so low volatility is persistent but can jump to high. In addition, from state 3 (moderate volatility) mostly stays moderate (89.5%), sometimes goes back to low volatility (7.4%), but can also jump to high volatility (3.1%), this stage is key transition point. Furthermore, from stage 4 (high volatility) highly persistent (95.1%), with tiny chances (1.2%, 3.7%) to slide to low or moderate volatility, so once volatility is high, it mostly maintains that level. In addition, from stage 5 (extremely high volatility) about 96.4% chance of persisting, but with a small chance (3.6%) to drop to extremely low volatility. In summary, the market mostly stays in the same volatility state dat-to-day, but moderate and extremely high volatility states are gateways where it can jump to other states faster.

Table 14 Emission Parameters

States	Emission means	Level	Emission std.v	Volatility level
S1	107.294	Extremely High-peak boom	11.269	Extremely high-moderate-high
S2	48.366	Low-moderate	5.337	Low-moderate
S3	82.004	High-elevated	3.647	High-lowest
S4	66.786	Moderate-balanced	4.491	Moderate-low
S5	157.663	Extremely low –lowest	23.476	Extremely low-highest

From the results obtained in Table 14 it is observed that from the emission means the spectrum captures a full cycle from crisis lows to expansion highs, with non-sequential state numbering (e.g., state 5 as lowest) highlighting model flexibility in regime assignment. It implies oil markets exhibit persistence in moderate-to-high states, aiding forecasts of stable pricing. Moreover, from the emission standard deviations it is observed that the parameters reveal asymmetric structure: extreme regimes (State 5 and 1) show elevated volatility (sigma > 10), signalling higher tail risks during booms/busts, while central states (3, 4, 2) have tighter spreads (sigma < 6) promoting predictability in normal conditions.

Table 15 Evaluation Metric Measures of HMM

MAE	MSE	RMSE
93.18904	10237.16	101.1788

The Table 15 consists of the evaluation metric measures that are used to assess the forecast power of Hidden Markov Model.

FIGARCH-5-State Hidden Markov Model

Table 16 Information Criterion of FIGARCH Models

Model	MAE	MSE	RMSE
FIGARCH (1, 1)-5-State HMM	4.337432	49.94298	7.067035
FIGARCH (1, 2)-5-State HMM	4.233755	48.68812	6.977687
FIGARCH (1, 4)-5-State HMM	4.331015	50.14674	7.081436
FIGARCH (1, 5)-5-State HMM	4.464736	51.01302	7.14234

From the results obtained in Table 16 it observed that FIGARCH (1, 2)-5-State HMM models outperforms the other models with least MAE, MSE and RMSE.

Table 17 Models Comparison

Model	MAE	MSE	RMSE
FIGARCH (1, 4)	4.286746	49.12025	7.008584
HMM	93.18904	10237.16	101.1788
5-States-FIGARCH (1, 2)-HMM	4.233756	48.68813	6.977688

From the results obtained in Table 17 it observed that the 5-States-FIGARCH (1, 2)-5-HMM outperforms the other models with least MAE, MSE and RMSE.

5-States-FIGARCH-HMM

Table 18 Initial State distribution (π)

Pr(extremely low)	Pr(low)	Pr(moderate)	Pr(high)	P(extremely high)
1	0	0	0	0

From the results obtained in Table 18 it is observed that since the value of probability value of extremely low volatility is 1, this indicates that the market is definitely at extremely low volatility state at time zero. This means we start by believing the oil market is (extremely low volatility) before a data is observed. The other states have zero chance initially. Hence, the market is in stable, moderate volatility phase before we observe any price changes.

Table 19 Transition Matrix (A)

States	To S1	To S2	To S3	To S4	To S5
From S1	0.743	0.257	0.000	0.000	0.000
From S2	0.000	0.641	0.112	0.076	0.171
From S3	0.000	0.018	0.926	0.008	0.055
From S4	0.054	0.146	0.000	0.799	0.000
From S5	0.000	0.050	0.015	0.029	0.905

From the results obtained in Table 19 it is observed that from state 1 (extremely low volatility) mostly stays in state 1 (74.3%), but sometimes jumps to low volatility (state 2) with 25.7%. Moreover, from state 2 (low volatility) it is observed that the volatility is very stable, staying 64.1% of the time, but small chance (11.2%, 7.6%, 17.1%) to jump suddenly to moderate, high, and extremely volatility (state 3, 4, and 5) respectively. In addition, from state 3 (moderate volatility) mostly stays moderate (92.6%), sometimes goes back to low volatility (1.8%), but can also jump to high volatility (8%) and extremely volatility (5.5%). Furthermore, from stage 4 (high volatility) highly persistent (79.9%), with tiny chances (5.4%, 14.6%) to slide to extremely low or low volatility. In addition, from stage 5 (extremely high volatility) about 90.5% chance of persisting, but with a small chance (5%, 1.5%, 2.9%) to drop to low, moderate or high volatility.

Table 20 Emission Parameters (B)

States	Emission means	Level	Emission std.v	Level
S1	-11.503	Extremely Low	24.093	Extremely low-highest
S2	-5.324	Low	3.984	Low-moderate
S3	-0.495	Moderate	1.750	Moderate-lowest
S4	9.454	High	4.759	Extremely high-higher
S5	1.300	Extremely High	3.555	High-moderate

From the results obtained in Table 20 it is observed that from the emission means this gradient (states 1 < 2 < 3 < 5 < 4) models a full volatility cycle, enabling regime-switching forecasts that highlight persistence in extremes for risk management in oils markets. Moreover, from the emission standard deviations low/moderate states show tighter spreads (less risk), while extremes (states1, and 4) exhibit amplified volatility, signalling higher tail risks during regime shifts. This refines forecasts by highlighting uncertainty in boom/bust phases.

Model Adequacy Checking

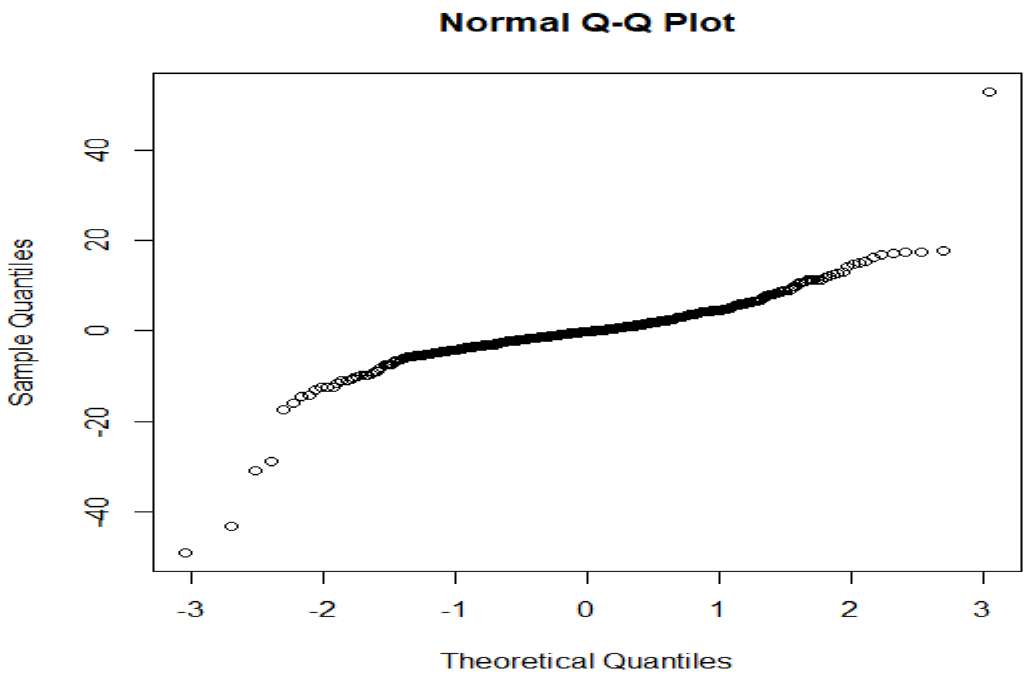


Figure 7 Residuals plots of FIGARCH (1, 2)-5-State HMM

From the above results obtained in Figure 7 it is observed that the residuals of the model have no much outliers and are approximately straight line. Thus, the errors are approximately normally distributed.

Table 21 Forecast Results

Month	Volatility forecast	Expected level
November	0.03050814	Likely to fall in moderate volatility
December	0.05410389	Likely to fall in moderate volatility
January	0.14353317	Likely to fall in moderate volatility
February	0.03214226	Likely to fall in moderate volatility
March	0.02828540	Likely to fall in moderate volatility
April	0.02474187	Likely to fall in moderate volatility

From the results obtained in Table 21 it is observed that the 5-States-FIGARCH (1, 2)-HMM forecasts moderate oils volatility (expected levels 2.50-2.58), with extremely high regime probabilities dropping from 3.05% in November to 0.02% in April. This indicates quick stabilization and low tail risks, favouring Low/Moderate states for predictable pricing and reduced hedging needs. The trend reflects regime persistence in calm conditions, supporting economic stability by early 2026.

From the time series plot it is observed that the time series exhibits trend with presence of structural breaks, this made the time series not stationary. Moreover, ADF and KPSS test confirmed that the time series is not stationary. The study revealed presence of long memory in the time series which was confirmed by GPH test. This long memory is persistent, therefore, is another behaviour that made the time series not stationary in addition to the trend behaviour observed. The study revealed excess of heteroscedasticity in the residuals of FIGARCH (1, 4) model which serves as the best model. The study modelled the varying level of the excess volatility observed in the residuals of the FIGARCH (1, 4) model through FIGARCH-5-States Hidden Markov Model, where the levels of the violating varying levels are categorized into 1 = extremely low, 2 = low, 3 = moderate, 4 = high, and 5 = extremely high. FIGARCH (1, 1)-5-Staes HMM was found as the best model with least MAE, MSE, and RMSE when compared with other models. The model revealed that the market is definitely at extremely low volatility state at time zero. Moreover, the model revealed that that from state 1 (extremely low volatility) mostly stays in state 1 (74.3%), but sometimes jumps to low volatility (state 2) with 25.7%. Moreover, from state 2 (low volatility) it is observed that the volatility is very stable, staying 64.1% of the time, but small chance (11.2%, 7.6%, 17.1%) to jump suddenly to moderate, high, and extremely volatility (state 3, 4, and 5) respectively. In addition, from state 3 (moderate volatility) mostly stays moderate (92.6%), sometimes goes back to low volatility (1.8%), but can also jump to high volatility (8%) and extremely volatility (5.5%). Furthermore, from stage 4 (high volatility) highly persistent (79.9%), with tiny chances (5.4%, 14.6%) to slide to extremely low or low volatility. In addition, from stage 5 (extremely high volatility) about 90.5% chance of persisting, but with a small chance (5%, 1.5%, 2.9%) to drop to low, moderate or high volatility. In addition, the model found that from the emission means the gradient (states 1 < 2 < 3 < 5 < 4) models a full volatility cycle, enabling regime-switching forecasts that highlight persistence in extremes for risk management in oils markets. Moreover, from the emission standard deviations low/moderate states show tighter spreads (less risk), while extremes (states1, and 4) exhibit amplified volatility, signaling higher tail risks during regime shifts. This refines forecasts by highlighting uncertainty in boom/bust phases. Furthermore, the study disclosed that the FIGARCH (1, 2)-5States HMM forecasts moderate oils volatility (expected levels 2.50-2.58), with extremely high regime probabilities dropping from 3.05% in November to 0.02% in April. This indicates quick stabilization and low tail risks, favouring Low/Moderate states for predictable pricing and reduced hedging needs. The trend reflects regime persistence in calm conditions, supporting economic stability by early 2026. In terms of model goodness of fit the FIGARCH (1, 2)-5States-HMM aligned with Chenxing and Qiao (2025), where their model outperforms other competing model.

The forecasts results obtained using the 5-Steates FIGARCH (1, 2)-HMM has the following implication:

The projected drop in extreme regimes suggests a shift toward predictable pricing by early 2026, reducing exposure to volatility spikes from geopolitical or supply disruptions. For crude oil, this implies lower hedging costs and more reliable supply chains, benefitting global trade amid current U.S energy policies. However, the initial 3.05% extremely high probability in November highlights residual risks from lingering factors like OPEC decisions, potentially amplifying short-term price swings. Overall, the forecast supports economic stability, with moderate inflation pressures on energy-dependent sectors like transportation and manufacturing.

CONCLUSION

The time series data exhibits a persistence trend over the study period. In addition, the time series shows fluctuation in both mean and variance, rendering it a non-stationary in levels. Furthermore, the study reveals the presence of long memory in the time series. Moreover, significant heteroscedasticity is observed in the residuals of FIGARCH (1, 2). The Five-States-FIGARCH (1, 2)-HMM was found to be the best performing model, yielding the lowest MAE, MSE, and RMSE when compared with the standalone HMM and FIGARCH models. The residuals of the model are approximately normally distributed indicating the model reduced volatility persistence. Finally the volatility forecast was made, which indicates that that the 5-States-FIGARCH (1, 2)-HMM forecasts moderate oils volatility (expected levels 2.50-2.58), with extremely high regime probabilities dropping from 3.05% in November to 0.02% in April. This indicates quick stabilization and low tail risks, favoring Low/Moderate states for predictable pricing and reduced hedging needs. The trend reflects regime persistence in calm conditions, supporting economic stability by early 2026.

REFERENCE

Alquraish , M., Abuhasel , K. A., Alqahtani, A. S., & Khadr, M. (2021). SPI-Based Hybrid Hidden Markov–GA, ARIMA–GA, and ARIMA–GA–ANN Models for Meteorological Drought Forecasting. Sustainability, 13(22), 12576.

Baum, L.E., & Petrie, T. (1986). Statistical Inference for Probabilistic Functions of Finite State Markov Chains. The Annals of Mathemeatical Statistics, 37(6), 1554-1563.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, 74(1), 3-30.

Balcilar, M., Ozdemir, Z. A., Ozdemir, H., & Wohar, M. E. (2020). Transmission of US and EU Economic Policy Uncertainty Shock to Asian Economies in Bad and Good Times. IZA Discussion Papers, 13274, Institute of Labor Economics.

Barroso, P., & Detzel, A. (2021). Do Limits to Arbitrage Explain the Benefits of Volatility Managed Portfolios? *Journal of Financial Econmics*, 140(3), 744-767.

Burhan, M., & Mohammed, A. (2024). Hybrid Model for Financial Forecasting using Maximal Overlap Discrete Wavelet Transform (MODWT), Demonstrating its Effectiveness in Capturing Volatility Patterns in Exchange Rates. *Jouurnal of Econometrics*, 10(3), 107-345.

Caldara, D., Lacoviello, M., Molligo, P., Prestipino, A., & Raffo, A. (2020). The Economic Effects of

Trade Policy Uncertainty. *Journal of Monetary Economics*, 109, 38-59.

Cross, J., & Nguyen, B. (2023). Oil Price Volatility and Supply Chain Distributions. Energy Economics, 108, 105-126.

Chenxing, L., & Qiao, Y. (2025). An Infinite Hidden Markov Model with GARCH for Short-Term Interest Rates. Online at https://mpra.ub.uni-muenchen.de/123200/ MPRA Paper No. 123200, posted 11 Jan 2025 14:25 UTC.

Deng, S., Xiang, Y., Nan, B., Tian, H., & Sun, Z. (2019). A Hybrid Model of Dynamic Time Wrapping and Hidden Markov Model for Forecasting and Trading in Crude Oil Market. Soft Computing (Springer), 23(12), 5679-5696.

Ellwanger, R., Jiang, L., & Sawatzky, B. (2023). Oil Price Shocks and the Macroeconomy: An Analysis of the Current Market Dynamics. Bank of Canada Staff Working Paper.

Hamilton, J. D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. Econometrica, 57(2), 357-284.

Hartwig, B., & Mahringer, S. (2023). Crude Oil Volatility : A Review of Recent Literature. Energy Economics, 106, 105-126.

Kilian, L., & Zhou, X. (2022). The Role of Global Supply and Demand in Oil Market Fluctuations. *Journal of International Economics*, 136, 103-126.

Mehrdoust, F., Noorani, I., & Kanniainen, J. (2024). Valuation of Option Price in Commodity Markets Described by a Markov-Switching Model: A Case Study of WTI Crude Oil Market, Mathematics and Computers in Simulation. *Journal of Econometrics*, 2(10), 345-678.

Sengupta, A., Das. A., and Guler, S. I. (2023). Hybrid hidden Markov LSTM for short-term traffic flow prediction.

Semmler, W., & Toure, M. (2024). Financial Fragility, Regime Change, and Monetary Policy in an Open Economy. Journal of International Economics, 10(23), 345-456.

Sharif, A., Aloui, C., & Yarovaya, L. (2020). Covid-19 Pandemic , Oil Prices, Stock Market, Geographical Risk and Policy Uncertainty Nexus in the US Economy: Fresh Evidence from the Wavelet-Based Approach. International Review of Financial Analysis, 70, 101476.

Tiwari, A. K., & Umar, Z. (2021). Existence of Long Memory in Crude Oil and Petroleum Products: Generalized Hurst Exponential Approach. Research in International Business and Finance. *Journal of Econometrics*, 89(23), 456-567.

Jung, Y., & Park, J. (2020). Scalable Hybrid HMM with Gaussian Process Emission for Sequential Time-series Data Clustering.

Wang, X., Shao, C., Yin, C., & Ji, X. (2018). Short-Term Traffic Flow Forecasting Method Based on ARIMA-GARCH Model. *Journal of Beijing Jiaotong University*, 42(4), 79. DOI: 10.11860/j.issn.1673-0291.2018.04.011.

Yin, L., Cao, H., & Xin, Y. (2024). Impact of Crude Oil Price Innovations on Global Stock Market Volatility: Evidence across Time and Space. International Review of Finnancial Analysis.

Zaki, J. FW., Ali-Eldin, A. M.T., Hussein, S. E., Saraya, S. F., & Areed, F. F. (2019). Time Aware Hybrid Hidden Markov Models for Traffic Congestion Prediction. *International Journal on Electrical Engineering and Informatics* - Volume 11, Number 1, March 2019.

Zhang, Y., Ma, F., & Wang, Y. (2022). Oil Price Volatility and Macroeconomic Fundamentals: A new Perpective. Energy Economics, 106, 105-126.