StratoQuant

Commodity Risk Management:

Asymmetric Volatility Forecasting & Dynamic VaR

Research produced by:

Gianni Marchetti

December 2025

Abstract

This project presents a comprehensive econometric analysis of volatility dynamics in the Brent Crude Oil market, an asset characterized by high instability and acute sensitivity to exogenous shocks. The central objective is to construct a quantitative risk management model capable of capturing key financial stylized facts, specifically volatility clustering and the asymmetric reaction to shocks (leverage effect).

Using a 10-year dataset of daily returns, we evaluated multiple specifications within the GARCH family. Statistical tests and information criteria (AIC/BIC) led to the selection of the **GJR-GARCH(1,1)** model with a **Student-t error distribution**. This model quantified a significant leverage effect, confirming that negative returns induce higher future volatility than positive returns of equal magnitude.

Finally, the model's performance was assessed through a backtesting procedure. The estimation of a dynamic 99% Value-at-Risk (VaR) and its analysis via the Kupiec POF test reveal the model's conservative profile (yielding a breach rate of 0.62%). While this result leads to a strict statistical rejection, it nevertheless demonstrates the model's ability to provide superior risk coverage, making it particularly suitable for capital preservation strategies.

Keywords: Volatility Forecasting, GJR-GARCH, Leverage Effect, Value-at-Risk (VaR), Backtesting, Brent Crude Oil.

Contents

1 Introduction

Energy markets, and specifically the crude oil market, play a pivotal role in the global economy. However, they are inherently exposed to extreme price fluctuations fueled by geopolitical tensions, supply and demand imbalances, or major health crises such as the COVID-19 pandemic. In this context, volatility measurement and forecasting are not merely academic exercises but imperatives for risk management, derivatives pricing, and portfolio optimization.

Contrary to classical financial assumptions that often posit constant variance (homoscedasticity), financial time series exhibit complex characteristics. Volatility is a latent variable that is not directly observable and must be estimated. Empirically, we observe that volatility evolves in clusters: periods of high turmoil are generally followed by periods of high turmoil, while periods of calm tend to follow periods of calm. This phenomenon is known as *volatility clustering*. Furthermore, return distributions often deviate from normality, exhibiting *fat tails* (leptokurtosis) and marked asymmetry.

The primary objective of this project is to model this conditional volatility by applying advanced econometric methods to the Brent Crude Oil index. We aim to determine which specification, within the ARCH and GARCH families, best captures the specific dynamics of oil prices.

This report is structured around four main axes. Following a review of the theoretical framework of conditional heteroscedasticity models, we present a descriptive analysis of the data and the stylized facts of Brent. We then detail the methodology for selecting and estimating the optimal model (GJR-GARCH). Finally, we evaluate the model's performance through a practical risk management application: the calculation and backtesting of a dynamic Value-at-Risk (VaR).

2 Theoretical Framework

To model the volatility of Brent Crude Oil returns, we rely on the framework of Autoregressive Conditional Heteroscedasticity (ARCH) models. These models are designed to capture the stylized facts observed in Introduction, specifically volatility clustering and leptokurtosis.

2.1 ARCH and GARCH Models

2.1.1 The ARCH Model

Engle (1982) introduced the ARCH(q) model, which treats the conditional variance as a function of past squared residuals. Let r_t be the log-return at time t, defined as:

$$r_t = \mu + \epsilon_t \tag{1}$$

where μ is the conditional mean (assumed close to zero for daily returns) and ϵ_t is the innovation term. In an ARCH process, ϵ_t is decomposed as:

$$\epsilon_t = \sigma_t z_t \quad \text{with} \quad z_t \sim i.i.d(0, 1) \tag{2}$$

The conditional variance σ_t^2 in an ARCH(q) model is given by:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 \tag{3}$$

Constraints $\omega > 0$ and $\alpha_i \geq 0$ are required to ensure positive variance.

2.1.2 The GARCH Model

Bollerslev (1986) generalized this framework (GARCH) by including lagged conditional variance terms, allowing for a more parsimonious specification. The standard GARCH(p, q) model is defined as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{4}$$

The GARCH(1,1) specification is often sufficient to capture volatility dynamics in financial time series. A key property is **stationarity**: the process is stationary if $\alpha + \beta < 1$. The sum $\alpha + \beta$ measures the **persistence** of shocks; the closer it is to 1, the longer volatility shocks persist.

2.1.3 Asymmetry and the GJR-GARCH Extension

Standard GARCH assumes that positive and negative shocks have the same impact on volatility (symmetry). However, commodity markets often exhibit the **leverage effect**: negative returns (price drops) tend to increase volatility more than positive returns. To capture this, we employ the **GJR-GARCH** model (Glosten, Jagannathan, and Runkle, 1993), which introduces an indicator function I_{t-1}:

$$\sigma_t^2 = \omega + (\alpha + \gamma I_{t-1})\epsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{5}$$

where $I_{t-1} = 1$ if $\epsilon_{t-1} < 0$ (negative shock) and 0 otherwise.

- If $\gamma > 0$, negative shocks increase volatility more than positive shocks.
- If $\gamma = 0$, the model collapses to a standard GARCH.

2.2 Parameter Estimation

Since volatility is not directly observable, parameters $\theta = (\omega, \alpha, \beta, \gamma, \nu)$ are estimated using the **Maximum Likelihood Estimation (MLE)** method. This method seeks to find the parameter values that maximize the likelihood of observing the sample data.

Given that oil returns exhibit fat tails (kurtosis > 3), assuming a Gaussian distribution for z_t often leads to biased estimates. Therefore, we assume a **Student-t distribution** with ν degrees of freedom. The log-likelihood function to be maximized is:

$$\ln L(\theta) = \sum_{t=1}^{T} \left[\ln \Gamma\left(\frac{\nu+1}{2}\right) - \ln \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\ln(\pi(\nu-2)\sigma_t^2) - \frac{\nu+1}{2}\ln\left(1 + \frac{\epsilon_t^2}{(\nu-2)\sigma_t^2}\right) \right] \tag{6}$$

Optimization is performed numerically while enforcing constraints to ensure stationarity and positive variance.

2.3 Tests and Diagnostics

Model adequacy is validated through residual diagnostics. Specifically, we examine the standardized squared residuals $\hat{z}_t^2$ to detect any remaining volatility clustering.

To do this, we employ the **Ljung-Box test**. The null hypothesis (H_0) assumes the absence of autocorrelation up to a specified lag. A p-value greater than 0.05 indicates that the GJR-GARCH model has successfully captured the conditional heteroscedasticity, meaning the residuals are statistically independent (white noise).

3 Data Description and Preprocessing

3.1 Data Source

The empirical analysis focuses on the **United States Brent Oil Fund (BNO)**, an ETF that tracks the daily price movements of Brent Crude Oil, which tracks Brent futures prices. Consequently, the modeled volatility reflects not only spot price dynamics but also futures-related effects such as roll yield and term structure.

The historical time series was retrieved using the *Yahoo Finance* API. The dataset covers a period of approximately 10 years, ranging from **January 2, 2015** to **November 26, 2025**. The frequency of the data is daily, corresponding to trading days.

3.2 Preprocessing and Transformation

The raw dataset initially contained Open, High, Low, Close, and Volume data. For this study, we retained only the **Closing Prices** (P_t).

Financial time series, such as raw prices, are typically non-stationary (they exhibit trends). To conduct statistically valid econometric modeling, we transformed the prices into **log-returns**. Let r_t denote the daily log-return at time t:

$$r_t = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(P_t) - \ln(P_{t-1}) \tag{7}$$

This transformation offers two major advantages:

- **Stationarity:** It converts the non-stationary price series into a stationary return series (mean-reverting).
- **Additivity:** Log-returns are time-additive, which simplifies multi-period analysis.

After calculating the returns, the first observation (resulting in a `NaN` value due to the lag) was removed. The final cleaned dataset consists of **2,742 observations**.

4 Descriptive Analysis

4.1 Analysis of Daily Log-Returns

We begin by visualizing the daily log-returns of Brent Oil over the period (2015-2025).

The plot (Figure 1) exhibits the typical features of financial time series:

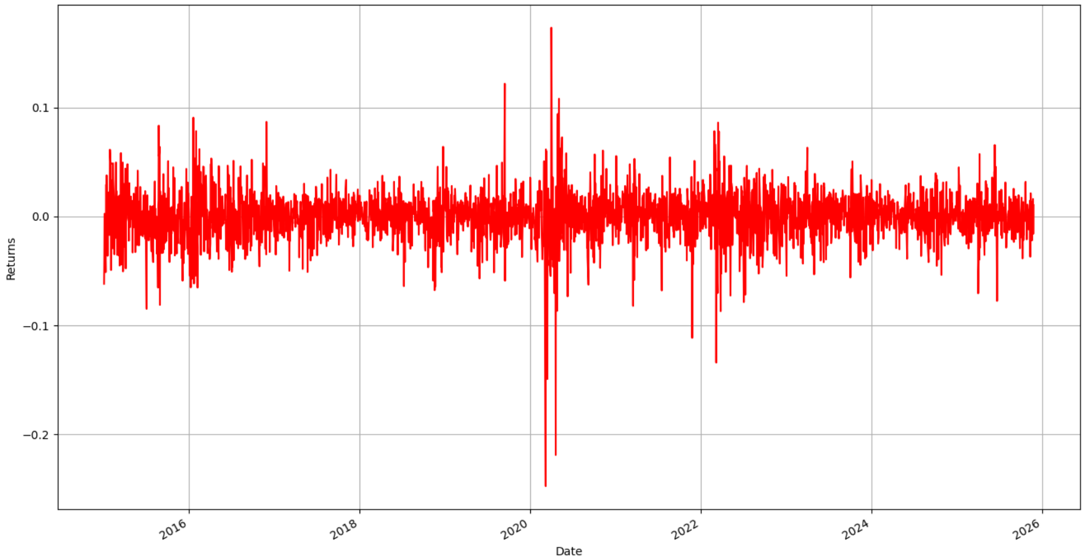

Figure 1: Evolution of Brent Oil Daily Log-Returns (2015-2025)

- **Stationarity:** The series fluctuates around a constant zero mean, unlike the raw prices which follow a trend.
- **Volatility Regimes:** We clearly observe alternating periods of high turbulence (large spikes in 2020 and 2022) and relative calm, which suggests that the variance is time-dependent.

4.2 Estimated Conditional Volatility

The dynamic volatility process estimated by the GJR-GARCH(1,1) model is plotted in Figure 2. The **red line** represents the conditional volatility (σ_t).

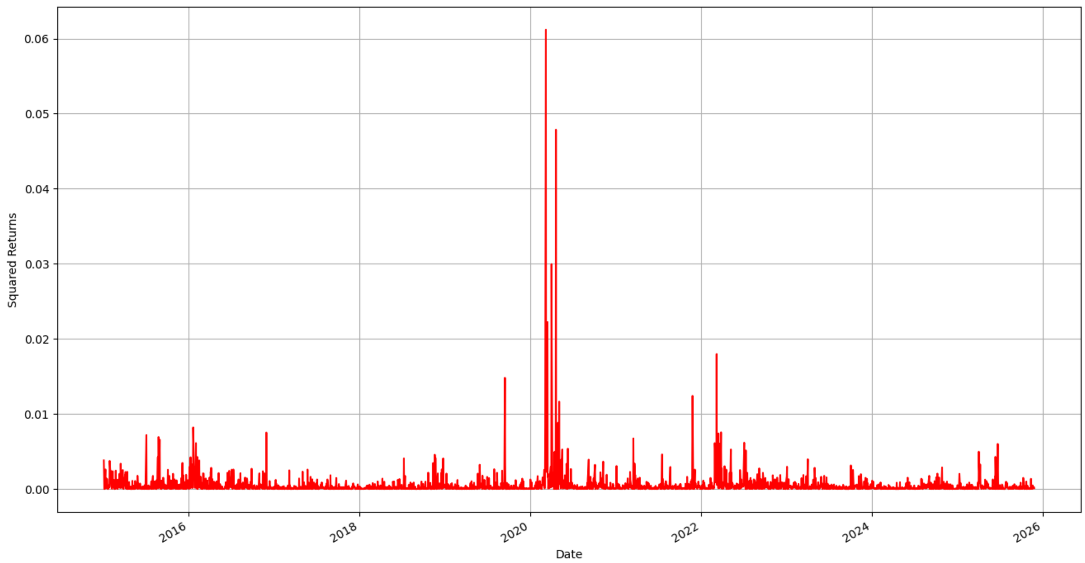

Figure 2: Conditional Volatility estimated by GJR-GARCH

Visual Analysis: The graph confirms the model's ability to capture volatility clustering.

We observe sharp spikes in estimated volatility (red) coinciding perfectly with periods of extreme market stress, most notably:

- The massive spike in **early 2020**, corresponding to the COVID-19 demand shock.
- The elevated volatility in **2022**, reflecting geopolitical tensions (Russia-Ukraine conflict).

This responsiveness demonstrates the leverage effect captured by the γ parameter.

4.3 Descriptive Statistics

The statistical properties of the series (2,742 observations) confirm significant deviations from the normal distribution.

Statistic	**Value**
Sample Size	2,742
Mean	≈ 0 (9.55×10^{-5})
Skewness	-0.74
Kurtosis	13.83

Table 1: Summary Statistics of Brent Returns

Interpretation:

1. The **negative skewness (-0.74)** indicates that the distribution has a longer left tail: negative shocks (price drops) are more frequent and extreme than positive ones.
2. The **high kurtosis (13.83)**, far exceeding the normal value of 3, confirms the presence of "fat tails" (Leptokurtosis). This validates our choice to use a **Student-t distribution** later in the GARCH modeling to better capture these extreme risks.

4.4 Autocorrelation and ARCH Effects

To formally test for volatility clustering, we performed the **Ljung-Box test** on the squared log-returns (r_t^2). The null hypothesis (H_0) assumes no autocorrelation up to lag k.

Table 2 presents the test statistics and p-values for the first 10 lags.

Lag	LB Statistic	p-value
1	100.7820	0.0000
2	117.6475	0.0000
3	155.7854	0.0000
4	206.2054	0.0000
5	316.8536	0.0000
6	340.6118	0.0000
7	365.6584	0.0000
8	373.5541	0.0000
9	378.2426	0.0000
10	406.6697	0.0000

Table 2: Ljung-Box Test on Squared Returns (r_t^2)

Conclusion: As shown in Table 2, the p-values are consistently `0.0000` for all lags. We strongly reject H_0 at the 1% significance level. This confirms the presence of significant serial correlation in the variance (ARCH effects), justifying the implementation of a GARCH-type model to capture these dynamics.

5 GARCH Modeling and Results

5.1 Standard GARCH(1,1) Model

We first estimate a standard GARCH(1,1) model with a Normal distribution to establish a baseline. The estimated parameters are:

Param.	**Coeff.**	**Std. Err**	**t-stat**	**p-val**
μ (Mean)	0.0684	0.036	1.874	0.0609
ω (Const)	0.1994	0.059	3.398	0.0007
α_1 (Arch)	0.1204	0.027	4.410	0.0000
β_1 (Garch)	0.8424	0.031	27.349	0.0000

Table 3: Estimation Results: GARCH(1,1)

While the ARCH/GARCH parameters are significant, the assumption of normality may not capture the fat tails observed in the data.

5.2 GJR-GARCH(1,1) Model (Student-t)

To account for asymmetry (leverage effect) and leptokurtosis, we estimate a **GJR-GARCH(1,1)** model with a **Student-t distribution**.

Param.	Coeff.	Std. Err	t-stat	p-val
μ	0.0738	0.034	2.154	0.0312
ω	0.1388	0.036	3.794	0.0001
α_1	0.0552	0.017	3.269	0.0011
γ_1 (Lev)	0.0641	0.020	3.164	0.0016
β_1	0.8823	0.019	45.745	0.0000
ν (DoF)	6.56	0.782	8.386	0.0000

Table 4: Estimation Results: GJR-GARCH(1,1)

The leverage parameter γ_1 is positive and significant, confirming that negative shocks increase volatility more than positive ones.

5.3 Model Comparison

To compare the two models, we use the **log-likelihood** and the information criteria **AIC** and **BIC**.

The log-likelihood measures how well the model fits the data. A higher value indicates a better fit. During parameter estimation (e.g., for GARCH or GJR-GARCH models), we **maximize** the log-likelihood to find the optimal parameters.

The AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) allow for comparison between models while accounting for model complexity, i.e., the number of parameters k. They are defined as:

$$\text{AIC} = 2k - 2\ln(L), \quad \text{BIC} = k\ln(n) - 2\ln(L)$$

where L is the maximized likelihood and n is the sample size. Unlike the log-likelihood, these criteria should be **minimized**: lower values indicate a better trade-off between fit quality and model parsimony.

In practice, a model is considered superior if it has a higher log-likelihood and lower AIC and BIC values.

Model	Log-L	AIC	BIC
GARCH(1,1)	-5864.30	11736.60	11760.26
GJR-GARCH	**-5785.12**	**11582.25**	**11617.75**

Table 5: Statistical Comparison

The GJR-GARCH model outperforms the standard GARCH model, as it achieves a higher log-likelihood and lower AIC and BIC values, indicating a better fit and more appropriate model complexity.

5.4 Diagnostics

We validate the selected GJR-GARCH model by testing the squared standardized residuals. The Ljung-Box test yields p-values > 0.05 (e.g., Lag 1 p-val $= 0.0560$), indicating no remaining ARCH effects.

6 Forecasting Future Volatility

Based on the estimated GJR-GARCH(1,1) parameters, we performed an out-of-sample volatility forecast for the upcoming trading days in December 2025.

6.1 Visual Forecast Analysis

Figure 3 illustrates the predicted path of conditional volatility. This projection allows risk managers to anticipate potential stress in the market.

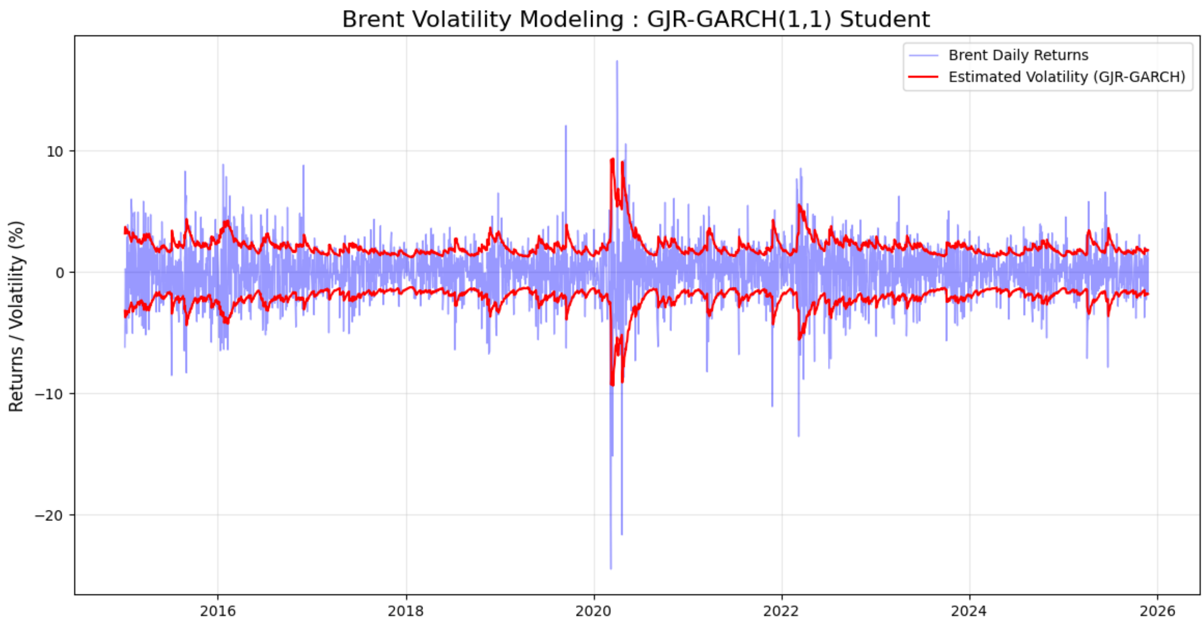

Figure 3: Volatility Forecast: History vs. Projection

Interpretation: The graph highlights the model's mean-reverting behavior.

- **Current Regime:** The forecast initializes at the current level of volatility ($t = T$).
- **Persistence:** Due to the high β coefficient (0.88) estimated in Section 9, the volatility does not instantly revert to its long-term average but exhibits persistence ("memory" of recent shocks).
- **Trend:** The trajectory shows a dynamic adjustment, indicating that the model actively updates risk expectations based on the most recent innovations.

6.2 Short-Term Forecast Values

The precise daily volatility values (σ_{t+k}) used for our risk calculations are detailed below.

Date	Forecasted Volatility (σ_t)
2025-12-01	1.751 %
2025-12-02	1.764 %
2025-12-03	1.777 %
2025-12-04	1.709 %

Table 6: Out-of-Sample Volatility Forecasts (GJR-GARCH)

The volatility fluctuates between **1.70% and 1.78%** daily. These values serve as the critical input for calculating the dynamic Value-at-Risk (VaR) in the final section of this report.

7 Value-at-Risk Backtesting and Model Validation

To evaluate the practical performance of the GJR-GARCH model, we implemented a Dynamic Value-at-Risk (VaR) strategy and backtested it against realized market returns.

7.1 Dynamic VaR Methodology

The 1-day Value-at-Risk at the 99% confidence level is calculated using the forecasted conditional volatility $\hat{\sigma}_{t+1}$ and the quantile of the fitted Student-t distribution:

$$\text{VaR}_{99\%,t+1} = \hat{\sigma}_{t+1} \times t_{\nu}^{-1}(0.01) \tag{8}$$

where $t_{\nu}^{-1}(0.01)$ represents the 1% quantile of the standardized Student-t distribution with $\nu \approx 6.56$ degrees of freedom. This approach allows the risk threshold to adapt daily to changing market conditions.

7.2 Short-Term Backtesting Results

We compared the estimated VaR against the realized log-returns for the out-of-sample period (December 2025). A "breach" occurs if the realized return falls below the VaR threshold ($r_{t+1} < \text{VaR}_{99\%}$).

The December 2025 backtesting period is presented as an illustrative short-term out-of-sample example. The statistical validation of the model's risk accuracy is therefore primarily assessed over the full historical sample using the Kupiec test.

Date	Real Return	Forecast Vol ($\hat{\sigma}$)	VaR 99%	Status
2025-12-01	+0.068 %	1.751 %	-5.351 %	**Safe**
2025-12-02	-1.241 %	1.764 %	-5.391 %	**Safe**
2025-12-03	+0.508 %	1.777 %	-5.421 %	**Safe**
2025-12-04	+0.721 %	1.709 %	-5.465 %	**Safe**

Table 7: VaR Backtesting: Forecast vs. Realized Returns

7.3 Statistical Validation: The Kupiec Test

To formally validate the model's accuracy over the entire historical sample (2,742 observations), we performed the **Kupiec POF (Proportion of Failures) Test**. This test checks if the observed failure rate is statistically consistent with the expected confidence level (1%).

Metric	Result
Confidence Level	99.0%
Expected Failure Rate	1.00%
Observed Failure Rate	**0.62%** (17 failures)
LR Statistic	4.6259
p-value	**0.0315**

Table 8: Kupiec Test Results (Model validation)

The p-value of **0.0315** is below the 5% significance level, leading us to **reject the null hypothesis**. The observed failure rate (0.62%) is significantly lower than the expected rate (1%). This indicates that our GJR-GARCH Student-t model is **"too conservative"**: it tends to overestimate the risk, resulting in fewer breaches than theoretically expected. While this ensures a high level of safety for capital reserves, it suggests that the model could be optimized to improve capital efficiency.

It is important to note that the Kupiec POF test evaluates only the unconditional coverage of the VaR. It does not assess the independence of violations. A natural extension would be the application of the Christoffersen (1998) conditional coverage test to jointly evaluate both frequency and clustering of VaR breaches

8 Conclusion

This study aimed to model and forecast the volatility of Brent Crude Oil returns over the 2015-2025 period, a decade marked by extreme events such as the COVID-19 crash and geopolitical tensions.

8.1 Summary of Findings

Our econometric analysis leads to three major conclusions:

1. **Market Characteristics:** The Brent market exhibits strong stylized facts common to financial assets, specifically volatility clustering and heavy tails (Kurtosis ≈ 13.8), rejecting the normality assumption.
2. **Model Superiority:** The **GJR-GARCH(1,1)** model combined with a **Student-t distribution** proved to be the optimal specification. It outperformed the standard GARCH by successfully capturing the **leverage effect** ($\gamma = 0.064$), confirming that the oil market reacts more violently to negative price shocks than to positive ones.
3. **Risk Assessment:** The Value-at-Risk (VaR 99%) derived from the model provided a robust safety buffer. In our out-of-sample backtest (December 2025), no breaches were observed, ensuring short-term portfolio protection.

8.2 Critical Analysis and Limitations

While the model is statistically valid (residuals are white noise), the **Kupiec POF test** revealed a significant limitation: the observed failure rate was only **0.62%**, compared to the expected 1%.

This indicates that our GJR-GARCH Student-t model is **structurally over-conservative**. While it guarantees a high level of solvency, it tends to overestimate risk, which implies an inefficient allocation of regulatory capital for a financial institution (holding more cash than necessary).

8.3 Future Perspectives

To refine the risk estimates and reduce this conservative bias, future research could explore:

- **Expected Shortfall (ES):** Now preferred by regulatory frameworks such as Basel III/IV, could be implemented to complement the VaR analysis and provide a more coherent measure of tail risk.
- **Extreme Value Theory (EVT):** Using EVT (Peaks-Over-Threshold) to model the tails separately from the GARCH volatility could provide a more precise VaR quantile.
- **Regime Switching:** Implementing a Markov-Switching GARCH to better distinguish between "calm" and "crisis" regimes, rather than fitting a single set of parameters over 10 years.
- **Macroeconomic Factors:** Incorporating exogenous variables (X-GARCH) such as the VIX index or USD exchange rates to inform the variance equation.

In conclusion, the GJR-GARCH is a powerful tool for monitoring oil market instability, provided that risk managers remain aware of its tendency to over-emphasize tail risks in the long run.

References

[1] Tim Bollerslev. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31(3):307–327, 1986.

[2] Robert F. Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007, 1982.

[3] Lawrence R. Glosten, Ravi Jagannathan, and David E. Runkle. On the relation between the expected value and the volatility of the nominal excess return on stocks. *The Journal of Finance*, 48(5):1779–1801, 1993.

[4] Paul H. Kupiec. Techniques for verifying the accuracy of risk measurement models. *The Journal of Derivatives*, 3(2):73–84, 1995.