Modeling Crude Oil Futures Volatility under Structural Breaks: An Evaluation of GARCH, MSGARCH, and Tree-Based Approaches

Course: Financial Volatility (Spring 2025)

Jacopo Rauseo (21-327-663)
Jana Cvetinovska (21-327-663)
University of St. Gallen

Date: May 22, 2025

Contents

1 Abstract

This paper compares different models for the prediction of crude oil futures price volatility, focusing on the period from 2018 to May 2025. We evaluate the performance of various GARCH-type models—standard GARCH, Markov-Switching GARCH (MSGARCH), and tree-structured GARCH—using West Texas Intermediate (WTI) crude oil futures data. Given the presence of structural breaks, particularly during the COVID-19 pandemic, subsequent recovery, and Russia-Ukraine war, we implement our models in a way that accounts for the change in regime. We further explore asymmetries in volatility dynamics by incorporating E-GARCH and GJR-GARCH models as well as using a Markov switching model and a tree-based regime-switching model based on macroeconomic covariates to capture heterogeneous volatility behavior. When estimated with a rolling window, the GARCH-family models (GARCH, EGARCH, GJRGARCH) generally outperform the MSGARCH and Tree-based models in forecasting volatility for a horizon of one month (20-25 trading days), especially when evaluated using MSE and QLIKE, with GJRGARCH ranking highest overall. However, improvements might be possible by using better proxies for volatility or more advanced tree-based techniques like ensembles.

2 Introduction

In recent years, the trajectory of oil prices has been marked by a high degree of uncertainty. A combination of economic instability and geopolitical crises has contributed to an exceptionally volatile environment in global commodity markets. The COVID-19 pandemic initially reduced economic activity and led to a sharp decline in oil prices, while Russia's invasion of Ukraine triggered a new upward cycle in global commodity prices. Figure 1 shows the evolution of oil futures in the last 7 years with some major events.

In this climate of uncertainty, effectively modeling energy prices plays an important role not only for portfolio and risk management but also for better understanding the impact on economic conditions. In fact, oil follows a non-trivial relationship with economic conditions: high oil prices can be the cause of economic slowdown (Auclert et al., 2023) as well its increase can be the consequence of a persistent period of economic growth, as in the 2002-2007 commodity cycle.

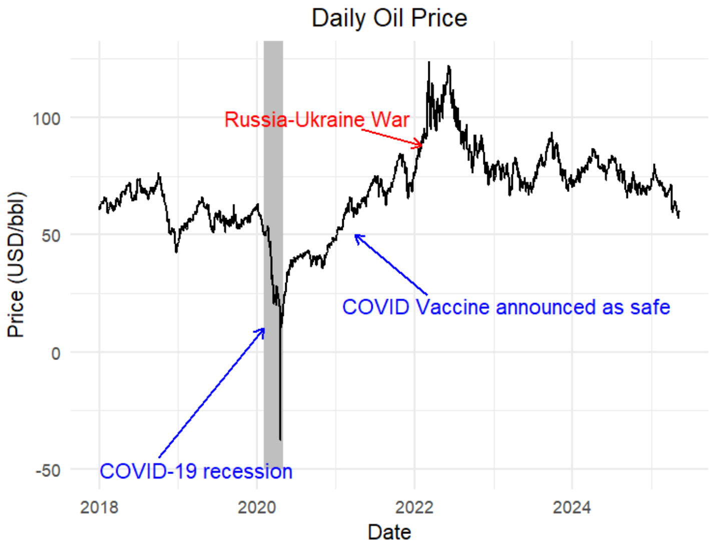

Figure 1: Oil Price (Source: Bloomberg)

In this paper, we conduct a direct comparison of various GARCH-type models of the forecast the volatility of crude oil in the period between 02.04.2025 and 7.05.2025. The objective of this comparison is twofold: (i) to replicate the key empirical properties of the two commodity price series, and (ii) to account for the structural breaks observed in the data.

3 Methods

In this section, we present our dataset and the general framework of our study, focusing on how we address the structural breaks present in the data. Considering that the time period under analysis includes at least two major events that significantly impacted financial and commodity markets, our approach involves (i) detecting the presence of these structural breaks and (ii) incorporating them appropriately into our models. In Subsection 3.2. we present our methodology to deal with structural breaks in a simple GARCH framework. We also briefly present two other GARCH-type models proposed in the literature to deal with multiple volatility regimes: the Markov-Swithcing GARCH (MSGARCH) and the tree-structured GARCH model.

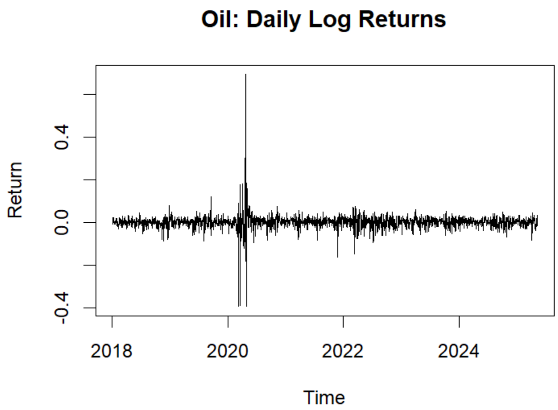

Figure 2: Oil Log returns

3.1 Dataset

Our dataset contains daily observations from January 2, 2018, to May 9, 2025, for the West Texas Intermediate (WTI) crude oil futures traded on the New York Mercantile Exchange. The data is taken from Bloomberg terminal.

Statistic	Log Return Oil
nobs	1850.00
Minimum	-0.39
Maximum	0.69
Mean	-0.0002
Median	0.001
Stdev	0.03
Skewness	2.07
Kurtosis	107.06
Jarque Bera	886913*

Table 1: Descriptive statistics for Log Return of Oil (* denotes rejection)

The main issue with the raw data is that the oil series shows a negative price for one trading day. This is quite exceptional and not very common for other financial instruments. To address this issue, we decided to fix the price of that observation to a value slightly greater than one (to make it possible to take the log after). Table 1 shows the main statistics of the log series. The kurtosis of the series is very high and the data is highly non-normal.

3.2 Dealing with structural breaks

In order to find structural breaks we search for breakpoints where the coefficients shift from one stable regression relationship to another. We search for m breakpoints that minimize the residual sum of squares (RSS) of the equation:

$$y_i = x_i^{\intercal}\beta_j + u_i \tag{1}$$

Where j denotes the index of each segment. We use the algorithm described by Bai & Perron (2003) for simultaneous estimation of multiple breakpoints which is based on a dynamic programming approach.

We first check for structural breaks in the daily oil log returns and find none. Then we look for breaks in the daily log returns squared.

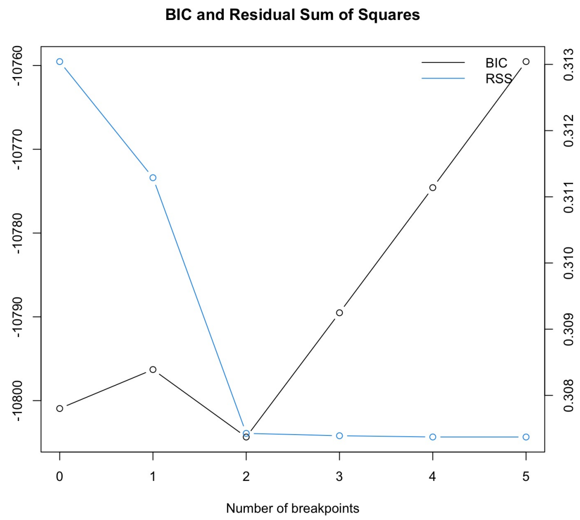

Figure 3: Optimal number of breaks

The RSS is minimized when two breakpoints are applied at February 21, 2020, and March 30, 2021, which correspond to the start of the COVID-19 pandemic and the announcement that the vaccine is safe and effective, respectively. We address these structural breaks by dividing the data into three segments at the identified points and then conducting tests and fitting GARCH models separately for each window.

The alternative method for checking the stability of the model is to use a rolling window approach. By estimating the GARCH-type of model in a rolling window we directly account for the change of regime. We also use the rolling window approach to compute the predictions of the three univariate GARCH models in section 4.

3.3 The general framework

Let r_t be the log return series. Our reference model is:

$$r_t = \mu_t(\cdot, \boldsymbol{\theta}) + a_t, a_t = \sigma_t(\cdot, \boldsymbol{\theta}) z_t \tag{2}$$

where the volatility σ_t, or the conditional variance σ^2, is modeled through different GARCH-type of models. The typical specifications apply and $z_t \overset{\text{iid}}{\sim} \mathcal{D}(0, 1)$. In our analysis we consider different functional forms of $\sigma_t^2 = h(a_{t-1}^2, \sigma_{t-1}^2, \boldsymbol{\theta})$.
In the simple framework presented in section 4.1. and 4.2., the estimation of the model is done in the $m + 1$ windows generated by the m structural breaks. We rewrite the GARCH

models using dummy variables. For example, the GARCH(1,1) equation becomes:

$$\sigma_t^2 = \sum_{j=0}^{m} D_{j,t}(\alpha_{j,0} + \alpha_{j,1}a_{t-1}^2 + \beta_{j,t-1}\sigma_{t-1}^2)$$

where $D_{j,t}$ is a dummy variable that takes 1 if $t \in [t_j; t_{j+1}]$ and 0 elsewhere.

3.4 A Markov-Switching GARCH Model (Ardia et al., 2019)

We extend our discussion by using a Markov Switching GARCH model (MSGARCH) to deal with the structural breaks in the data. The theoretical features of the model presented here and the R implementation are done according to the work of Ardia et al. (2019). Let a_t be a weak white noise process satisfying:

$$a_t = z_t \sigma_{k,t} \tag{3}$$

Then according the MRS-GARCH specification the process a_t follows:

$$a_t | (s_t = k, \mathcal{F}_{t-1}) \sim \mathcal{D}(0, h_{k,t}, \boldsymbol{\xi}_k) \tag{4}$$

where $\mathcal{F}_{t-1}$ denotes the information set, s_t is an integer-valued latent variable in $\{1, ...K\}$ representing the regime of the Data Generating Process (Reher and Wilfling, 2011), $h_{k,t}$ is the variance at time t, and the vector $\boldsymbol{\xi}_k$ indicates the shape parameters for the different GARCH model in the state k. The standardize innovation are:

$$z_{k,t} = \frac{a_t}{\sqrt{h_{k,t}}} \overset{\text{iid}}{\sim} \mathcal{D}(0, 1, \boldsymbol{\xi}_k) \tag{5}$$

Then we assume that the latent process s_t follows a first-order Markov chain with the following probability transition matrix P:

$$P = \begin{bmatrix} p_{11} & \cdots & p_{1,K} \\ \vdots & \ddots & \vdots \\ p_{K,1} & \cdots & p_{K,K} \end{bmatrix}$$

where the $p_{i,j}$ element represent the probability to switch from the state $s_{t-1} = i$ to $s_t = j$. The conditional variance follows a GARCH-type of model where the specifications of the shape parameters ensure the positiveness.

3.5 A Tree-structured GARCH model

We replicate the methodology introduced by Audrino, F., & Bühlmann, P. (2001). By We construct a binary tree that recursively splits the dataset based on seven covariates (Gold price, S&P500, Dollar Index, US governement bond 10Y, S&P500 Energy, Bloomberg Commodity index and the Caldara-Iacovello Geopolitical Risk index), aiming to partition the data into subsets where the volatility dynamics differ significantly (different regimes). The splitting is done by trying thresholds for each covariate, fitting the GARCH model at each step, and splitting if the split lowers the AIC, otherwise the node becomes a terminal node. After the split, the algorithm is called recursively on each subtree. We initially construct a large tree with the only stopping criterion being a minimum terminal node size of 300 in order to ensure enough observations to fit the GARCH on. Then, we proceed with a pruning step where if the split does not produce a significant decrease in AIC, it is removed.

We have chosen to fit a GARCH(1,1) model in the leaves since the tree structure should already take into account asymmetries, and a more complex model would worsen the overfitting risk that tree algorithms are already prone to.

Algorithm 1 Tree-Structured GARCH

Require: Dataset D, covariates X, return series r, minimum node size $n_{\min}$

function TreeGARCH(D)
    if $|D| < n_{\min}$ **then**
        return FitGARCH(r_D)
    end if
    $best\_aic \leftarrow \infty$
    $best\_split \leftarrow$ None
    for all covariates $x \in X$ **do**
        for all thresholds t for x **do**
            $D_{\text{right}} = \{x \in D : x \leq t\}$
            $D_{\text{right}} = \{x \in D : x > t\}$
            if $|D_{\text{left}}| < n_{\min}$ or $|D_{\text{right}}| < n_{\min}$ **then**
                continue
            end if
            FitGARCH(D_{left}) and FitGARCH(D_{left})
            Compute total AIC after split
            if total AIC < $best\_aic$ **then**
                Update $best\_aic$
                $best\_split = list(x, t)$
            end if
        end for
    end for
    if $best\_split =$ None **then**
        return FitGARCH(r_D)
    else
        return node with:
          left = TreeGARCH(D_{left})
          right = TreeGARCH(D_{right})
    end if
end function

4 Results

We begin by discussing the simple framework where we fit different GARCH-type of model to the three window that we define in section 3. The regression (1) for the mean of the process did not show any structural break, suggesting that the conditional mean parameters remain the same across the whole period of analysis. For a better identification of the order of the mean process we start by conducting an ARCH test up to lag 10, whose results are displayed in Table 2.

Window	LM Test	Q(5)	Q(10)
1	19.61 (0.03)	8.23 (0.14)	10.41 (0.40)
2	27.62 (0.002)	3.06 (0.68)	11.67 (0.30)
3	53.27 (0)	5.83 (0.32)	11.67 (0.49)

Table 2: ARCH Test and Corrected Box-Ljung Test
(p-values under brackets)

We reject the LM test at the 5% confidence level for the three windows, implying that there is statistical evidence of ARCH effects present in all sections. Therefore, we correct the confidence bounds of the ACF plot, shown in Figure 4. By performing the corrected Ljung box test for ARCH effect and looking at the ACF plot, we cannot reject the null hypothesis of zero correlation. Therefore we can consider that the process r_t is white noise and we don't have an ARMA form for the conditional mean. The fact that the three windows share the same mean process is consistent with the previous structural breaks analysis of the log return series.

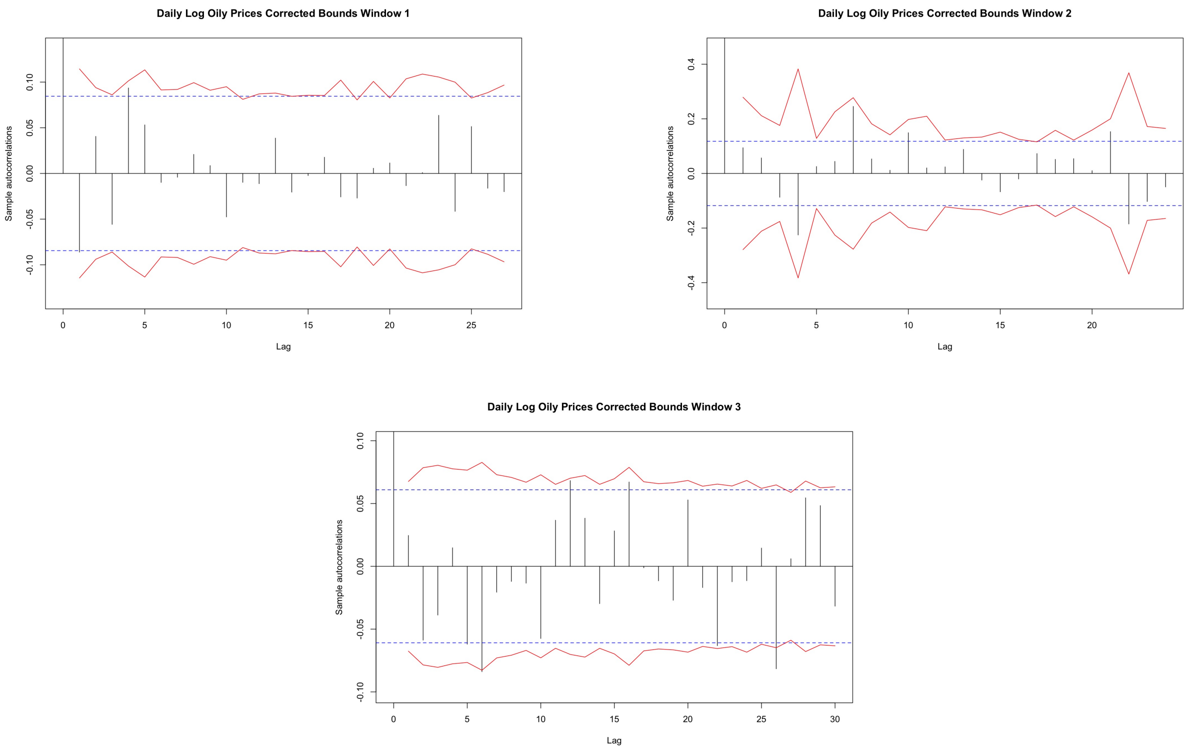

Figure 4: ACF of Daily Log Returns with Corrected Confidence Bounds

4.1 GARCH

We now estimate different types of GARCH(p,q) models and select the order using the Akaike information criteria that better fit the data in the three windows [1]

[1] Here the model are all estimated using student t distributed errors. We performed the same exercise using $z_t \stackrel{\text{iid}}{\sim} N(0,1)$, but the residuals were non-normal suggesting fat tails.

Model	**Window 1 (AIC)**	**Window 2 (AIC)**	**Window 3 (AIC)**
ARCH(1)	-5.090	-3.772	-4.722
ARCH(2)	-5.090	-3.772	-4.722
GARCH(1,1)	-5.110	-3.942	-4.792
GARCH(1,2)	-5.107	-3.942	-4.790
GARCH(2,1)	-5.107	-3.935	-4.791
GARCH(2,2)	-5.107	-3.935	-4.791
GARCH(2,3)	-5.101	-3.929	-4.790
GARCH(3,2)	-5.100	-3.931	-4.788
GARCH(3,3)	-5.097	-3.922	-4.788

Table 3: AIC values for the the three windows

According to the AIC criteria we recorded in Table 3 the best model fit is the GARCH(1,1) for windows 1 and 3, while GARCH(1,2) fits slightly better for window 2 [2]. From the behavior of the log return series we could have deducted that a GARCH model is better than a ARCH model, given the clear volatility clustering in the data, especially during the beginning of the COVID-pandemic.

Table 7 shows the coefficients for estimated by QML for the three GARCH models, as well as the conditions for the existence of the second-order statinarity $m_1 < 1$ and for the existence of the fourth moment $m_2 < 1$. Here we present the results for which we used the student t distribution for the shocks.

Parameter	Window 1	Window 2	Window 3
α_1	0.0511	0.3132	0.0797
β_1	0.9084	0.0691	0.8929
β_2	—	0.598	—
m_1	0.9595	0.980	0.9726
m_2	—	—	1.50

Table 4: GARCH(1,1) Parameter Estimates

The fourth moment of the GARCH processes doesn't exist. In addition, only for window 3 the fourth moment of the student t-distribution used for the shock exists. In the other two windows the degrees of freedom are $2 < n < 4$, implying that the t distribution has infinite variance. Finally, a visual check hints that the standardized residuals of the models and their squared counterparts show no sign of autocorrelation. Figure 5 shows the in-sample volatility estimated in the three windows.

[2]We used the rugarch package to estimate the GARCH models. The package divides the AIC by the sample size N (Ghalanos, 2024).

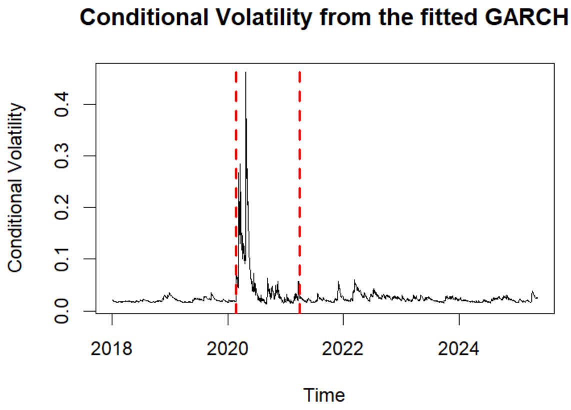

Figure 5: Estimated in-sample conditional volatility

4.2 Accounting for Leverage Effect

According to the literature, the existence of a leverage effect in the oil market remains ambiguous (Matar, 2013). While it is well-documented that, for many stocks, negative shocks tend to have a greater impact on volatility than positive shocks of the same magnitude, the evidence for oil prices is less conclusive. In some periods, price declines are associated with high volatility, consistent with typical negative leverage effects. However, in other cases, price increases — particularly following adverse news — can also lead to increased volatility, suggesting that asymmetric responses may vary depending on the underlying market conditions. Table 5 shows the correlation $\rho(r_t^+, r_{t-h})$.

h = 1	**h = 2**	**h = 3**	**h = 5**	**h = 10**	**h = 20**
0.044	-0.004	-0.059	-0.025	0.074	-0.033

Table 5: Values for $\rho(r_t^+, r_{t-h})$ for selected lags

To assess the presence of leverage effect we can run a Negative Size Bias Test, whose results are shown in Table 6.

Window	**Negative Size Bias Test**	**Positive Size Bias Test**	**Sign Test**
1	-4.637e-02*	3.826e-02*	2.288e-04*
2	-0.2946211*	0.484059	-9.416e-05
3	-6.154e-02*	1.824e-02*	3.335e-04*

Table 6: Bias tests for Leverage Effects
** denotes significance at the 1% level*

The regression coefficient for the negative size bias test is significant at the 1% significance level for the three windows. To account for asymmetric behavior, we use GARCH estimate the volatility using EGARCH and GJR-GARCH with z_t student-t distributed. The in-sample fit of the E-GARCH and GJR-GARCH is shown in Figure 6 below. The parameters are computed separately for each of the three windows and the red dashed line indicates the structural breaks. The behavior of the estimated volatility is similar to the plain vanilla GARCH estimation.

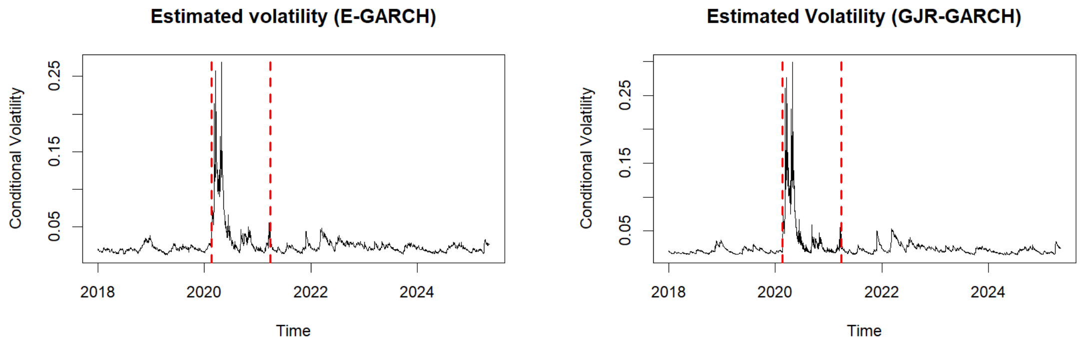

Figure 6: In-Sample volatility of asymmetric GARCH

Comparing the AIC for the three GARCH models we see that while in Window 1 an asymmetric model might offer a better fit, in windows 2 and 3 neither E-GARCH nor GJR-GARCH offer a substantial improvement in terms of goodness of fit.

Parameter	Window 1	Window 2	Window 3
GARCH	-5.110	-3.941	-4.791
EGARCH	-5.125	-3.927	-4.780
GJRGARCH	-5.120	-3.939	-4.790

Table 7: Summary of the AIC information criteria

The approach of separating the data into three fixed windows defined by structural breaks is overly rigid, and since the regime change likely does not occur in one day, using an approach that is more adaptive to smooth changes might be more appropriate. For this reason, from this point onward, except for the MSGARCH and the tree, we proceed with a rolling window approach to fitting our models and making predictions.

We use here a rolling window of size m = 1000 as suggested by Audrino and Chassot (2024) and we check for parameter stability. Longer rolling window sizes tend to yield smoother rolling window estimates than shorter sizes. The number of increments between successive rolling windows is 1 period, we partition the entire data set into N = T − m + 1 subsamples. The first rolling window contains observations for period 1 through m, the second rolling window contains observations for period 2 through m + 1, and so on. We then estimate the model using each rolling window subsamples and plot each estimate and point-wise confidence intervals over the rolling window index to see how the estimate changes with time.

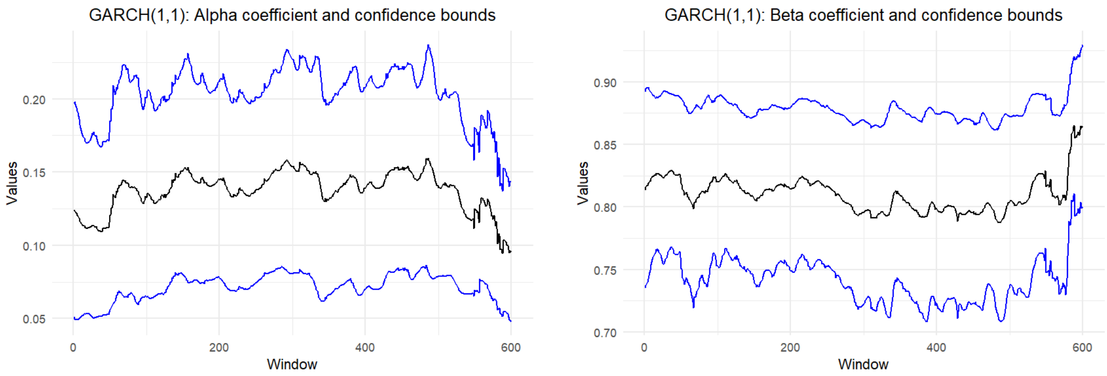

Figure 7: Stability of the model

4.3 MS-GARCH

We estimate a MSGARCH for the period between January, 1, 2018 to April, 31, 2024. We use a GJRGARCH as model for the different regimes and we start by comparing a MSGARCH with three regimes with one with only two. The MSGARCH with only two regimes yields a AIC of -8689.09, while the other one -8676.9. Therefore the MSGARCH with only two regimes fit better the data. From the estimated transition matrix, the probability to switch from state $s_{t-1} = 1$ to state $s_t = 2$ is 0.03, while the probability to switch from state $s_{t-1} = 2$ to $s_t = 2$ is 0.10. Figure 8 shows the in-sample volatility.

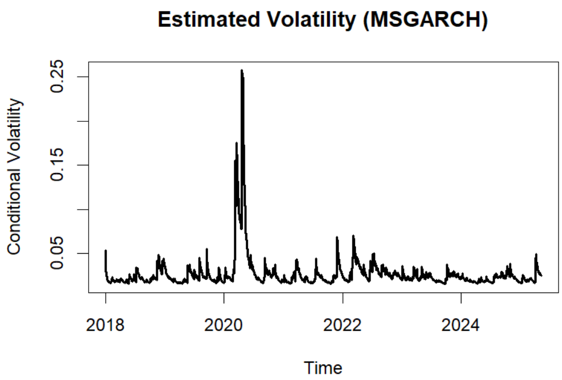

Figure 8: Estimated in-sample conditional volatility

4.4 Tree-GARCH

Figure 8 shows the resulting tree structured GARCH after pruning. Reported are the AICs normalized by the sample size for consistent comparison with the output of the fGarch package in R. The AIC for the overall model is -8676.57.

We have four splits that result in 5 terminal nodes where in terms of in-sample analysis, the AIC appears to be an improvement over the regular GARCH and asymmetric GARCH models.

Regular decision trees like CART are prone to overfitting, and this type of tree-structured GARCH might be even more vulnerable to this issue. We have taken some steps to mitigate overfitting by making splits based on AIC to avoid excessive complexity, pruning the tree based on likelihood, and imposing a minimum node size.

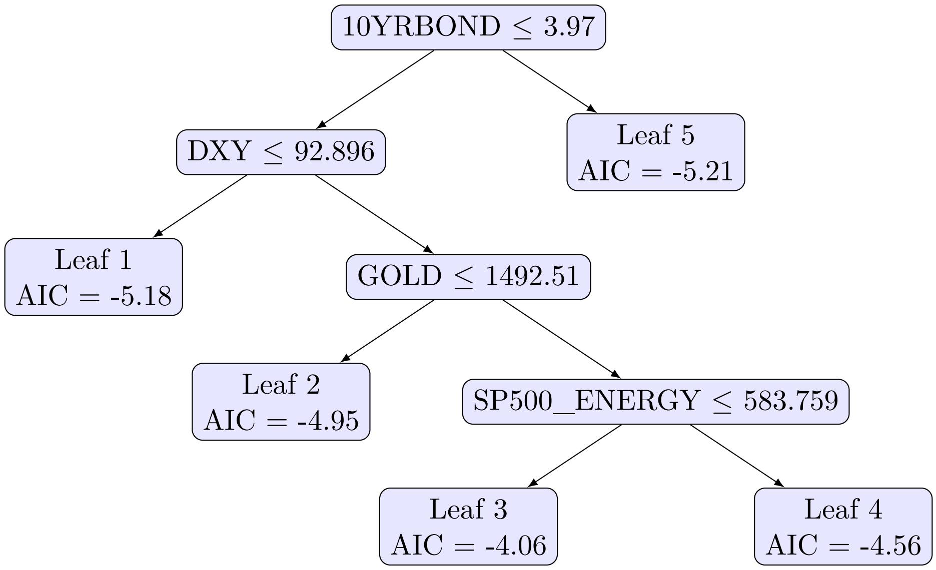

Figure 9: Pruned Decision Tree for GARCH Model

From the table below we can see that there is second-order stationarity in all leaves except leaf 2, and that the fourth moment exists only in leaves 1, 4 and 5.

Parameter	**Leaf 1**	**Leaf 2**	**Leaf 3**	**Leaf 4**	**Leaf 5**
ω	0.000072	0.000024	0.000092	0.000070	0.000012
α_1	0.150940	0.085546	0.213250	0.045702	0.060864
β_1	0.653550	0.915950	0.757230	0.843590	0.901520
m_2	0.804490	1.001496	0.97048	0.889292	0.962384
m_4	0.909587	-	-	0.801805	0.945631
Shape	6.1922	2.6947	3.386	10	10

Table 8: GARCH(1,1) model parameters in each leaf

The conditional volatility computed by the tree appears to have higher spikes than the asymmetric GARCH and Markov switching model, but lower spikes than the regular GARCH model.

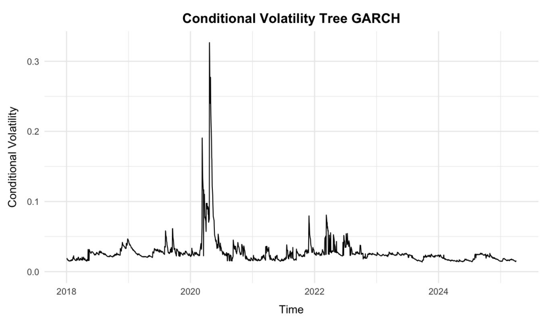

Figure 10: Estimated in-sample conditional volatility

5 Discussion

In this section, we conduct an out-of-sample evaluation of our models using a direct comparison approach. We set aside a forecast window of 25 periods, spanning from April 2, 2025, to May 7, 2025. For the MSGARCH and Tree-based models, we directly predict the volatility over this out-of-sample window. However, for the three GARCH models, such a long forecast horizon is less suitable due to their structure. Therefore, we adopt a rolling window approach with a window size of $m = 1000$ days, generating one-step-ahead forecasts that are iteratively combined to construct the $h = 25$ step forecast. As robust proxy for the volatility we use the squared returns r_t^2. In table 9 we summarize the main loss function for the different predictions.

	GARCH	EGARCH	GJRGARCH	MRSGARCH	TREE
MSE	2.32e-06	2.41e-06	2.29e-06	3.03e-06	2.92e-06
MAE	8.48e-04	8.44e-04	8.46e-04	8.55e-04	8.84e-04
QLIKE	-6.047	-6.020	-6.056	-5.187	-5.390

Table 9: Forecast Evaluation Metrics

To compare multiple models we use the Hansen's test of superior predictive ability (SPA). The null hypothesis of the test is that no other model forecast outperforms the chosen one. Table 10 shows the p-values for the SPA test.

	GARCH	EGARCH	GJRGARCH	MRSGARCH	TREE
MSE	0.485	0.514	0.838	0.066	0.062
MAE	0.760	0.730	0.748	0.397	0.761
QLIKE	0.532	0.516	0.532	0.655	0.520

Table 10: p-values for the Test for Superior Predictive Ability

Given the Mean Absolute Error (MAE) is not a robust statistic, we only consider the MSE and the QLIKE. It appears that the results of the SPA test do not indicate there is a statistically significant improvement that any model offers over another. We proceed by conducting a Model Confidence Set (MCS) test which operates by iteratively excluding models which underperform until we obtain a superior set of models which are indistinguishable from the best-performing one. We have chosen to conduct the test based on the QLIKE and MSE loss. The results of the MCS test with the QLIKE loss are given in the table below.

Model	**Rank**	**MCS**	**QLIKE**
GARCH(1,1)	2	1.0000	-6.047
EGARCH(1,1)	1	1.0000	-6.020
GJRGARCH(1,1)	3	1.0000	-6.056
MSGARCH	5	0.1326	-5.133
Tree-GARCH	4	0.4104	-5.662

Table 11: Model Confidence Set (MCS) results based on QLIKE

According to these results, MSGARCH and Tree-GARCH are excluded. Therefore, we can finally conclude that the prediction of the three GARCH models estimated with the rolling-window approach perform equally well.

In conclusion, we briefly discuss the limitations of our analysis. Contrary to what is often found in the literature, such as in Audrino and Bühlmann (2001) and Haas et al. (2004), we found that the MSGARCH and Tree-GARCH models perform worse than simple GARCH models. We propose several reasons for this outcome: (i) the volatility regimes are not clearly defined or segmented; aside from the brief "COVID shock" (Window 2 in Section 3.2), other periods appear to exhibit lower levels of volatility[3], (ii) potential overfitting, (iii) a forecast horizon that may be too long, and (iv) the rolling window approach appears to offer an advantage over the two "static" models. Although computationally intensive, a rolling window implementation might yield improved results for both the MSGARCH and Tree-GARCH models. Regarding the tree, there are some clear signs of overfitting. Our results could likely be improved with more data, splitting based on different covariates, or using ensemble models like random forest or gradient boosted trees instead of a single decision tree.

Finally, the volatility proxy used to calculate the loss functions is the square of the log returns. We might obtain significant results if we instead use the realized volatility as the proxy for evaluation, however we would need to obtain high-frequency data for that.

[3]It is worth noting that the war in Ukraine did not generate as much volatility in the oil market as it did in the gas market.

6 References

Ardia, D., Bluteau, K., Boudt, K., Catania, L., & Trottier, D.-A. (2019). Markov-Switching GARCH Models in R: The MSGARCH Package. *Journal of Statistical Software, 91*(4), 1–38. Retrieved from `http://www.jstatsoft.org/v91/i04/`

Auclert, A., Monnery, H., Rognlie, M., & Straub, L. (2022). Managing an Energy Shock: Fiscal and Monetary Policy. *Working Paper*. Retrieved from `https://www.dropbox.com/s/9qp4i6hdojdu3bt/AMS_EnergyShock.pdf`

Audrino, F., & Bühlmann, P. (2001). *Tree-structured Generalized Autoregressive Conditional Heteroscedastic Models*. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(4), 727–744. Retrieved from https://doi.org/10.1111/1467-9868.00309

Audrino, F., & Chassot, J. (2024). *Hard to Beat: The Overlooked Impact of Rolling Windows in the Era of Machine Learning*. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. Retrieved from https://papers.ssrn.com/abstract=5026062

Audrino, F., & Trojani, F. (2008). A General Multivariate Threshold GARCH Model with Dynamic Conditional Correlations. *Econometric Reviews, 27*(5), 497–522. `https://doi.org/10.1080/07474930801959610`

Bai, J., & Perron, P. (2002). *Computation and Analysis of Multiple Structural Change Models*. Journal of Applied Econometrics, 18(1), 1–22.

Ghalanos, A. (2024). *Introduction to the rugarch package (Version 1.4-3)*. Retrieved from https://cran.r-project.org/web/packages/rugarch/index.html

Haas, M. (2004). A New Approach to Markov-Switching GARCH Models. *Journal of Financial Econometrics, 2*(4), 493–530. `https://doi.org/10.1093/jjfinec/nbh021`

Hansen, P. R. (2005). *A Test for Superior Predictive Ability*. Journal of Business Economic Statistics, 23(4), 365–380.

Matar, W., Al-Fattah, S. M., Atallah, T., & Pierru, A. (2013). An Introduction to Oil Market Volatility Analysis. *OPEC Energy Review, 37*(3), 247–269. `https://doi.org/10.1111/opec.12017`

Reher, G., & Wilfling, B. (2015). Markov-Switching GARCH Models in Finance: A Unifying Framework with an Application to the German Stock Market. *Studies in Nonlinear Dynamics & Econometrics, 19*(1), 91–114. `https://doi.org/10.1515/snde-2014-0052`

Tsay, R. S. (2010). *Analysis of Financial Time Series* (3rd ed.). Wiley Series in Probability and Statistics. John Wiley & Sons.

7 Auxiliary Tools

We used ChatGPT to help us format and menage the figures and tables in the LaTeX file. We also used it for debugging purposes in the R code.

8 Declaration of Authorship

I hereby declare,

- that I have written this essay independently,
- that I have written the essay using only the aids specified in the index;
- that all parts of the essay produced with the help of aids have been declared;
- that I have handled both input and output responsibly when using AI. I confirm that I have therefore only read in public data or data released with consent and that I have checked, declared and comprehensibly referenced all results and/or other forms of AI assistance in the required form and that I am aware that I am responsible if incorrect content, violations of dataprotection law, copyright law or scientific misconduct (e.g. plagiarism) have also occurred unintentionally;
- that I have mentioned all sources used and cited them correctly according to established academic citation rules;
- that I have acquired all immaterial rights to any materials I may have used, such as images or graphics, or that these materials were created by me;
- that the topic, the thesis or parts of it have not already been the object of any work or examination of another course, unless this has been expressly agreed with the faculty member in advance and is stated as such in the thesis;
- that I am aware of the legal provisions regarding the publication and dissemination of parts or the entire thesis and that I comply with them accordingly;
- that I am aware that my essay can be electronically checked for plagiarism and for third-party authorship of human or technical origin and that I hereby grant the University of St. Gallen the copyright according to the Examination Regulations as far as it is necessary for the administrative actions;
- that I am aware that the University will prosecute a violation of this Declaration of Authorship and that disciplinary as well as criminal consequences may result, which may lead to expulsion from the University or to the withdrawal of my title.

By submitting this essay, I confirm through my conclusive action that I am submitting the Declaration of Authorship, that I have read and understood it, and that it is true.

Date and Signature
22.05.2025, JR and JC