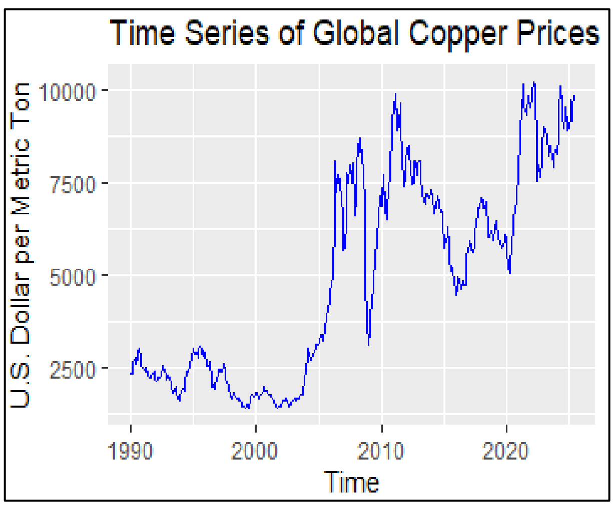
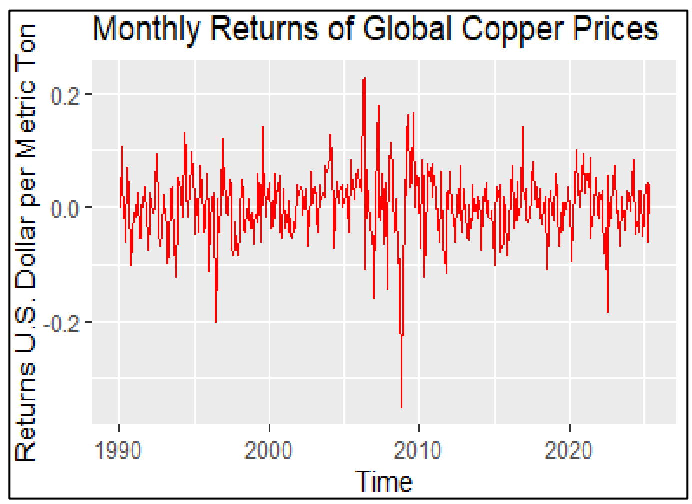
ARIMA/GARCH Modelling of Global Copper Prices Volatility

Background of the data

Copper is often referred to as “Dr. Copper” because it has a unique ability to gauge the overall health of the global economy. It’s a versatile metal used extensively in construction, electrical wiring, transportation, and manufacturing. This means that copper demand is closely tied to economic activity. The supply and demand of copper also affect its price, with a shortage leading to a price increase and the abundance of copper leading to a price drop. Political factors and speculative trading also affect the price of copper. The price of copper can fluctuate significantly, making investments in copper very risky. The values represent the benchmark prices, which are representative of the global market. Prices are period averages in nominal U.S. dollars. Our data is collected between 01 January 1990 and 01 June 2025 and recorded monthly. There are 426 observations. The data is taken from [Global price of Copper (PCOPPUSDM) | FRED | St. Louis Fed](), and a screenshot of the data is provided on Figure 1.

Table 1: a) Time Series plot of the data and b) Monthly Returns of the data

a)	b)
	

Copper price time series plot shows a gradual increase, being more drastic from 2005 due to China’s large copper demand for its bold industrialization. Around 2008, we see a large drop in copper prices to 3000 USD, due to the global financial crisis that led to a global recession. The prices then grew exponentially to staggering highs of 9000 USD, due to China’s resilient demand. Post Covid, 2020, Copper prices grew due to global demand for renewable energy sources.

The returns of copper show heavy volatility clustering, with evidence of an outlier.

	A	B
1	observation_date	PCOPPUSDM
2	1990-01-01	2365.55699088135
3	1990-02-01	2358.94340000000
4	1990-03-01	2625.70272275848
5	1990-04-01	2685.22689088135
6	1990-05-01	2740.34239088135
7	1990-06-01	2583.81490911865
8	1990-07-01	2769.00302275848
9	1990-08-01	2956.39555455933
10	1990-09-01	3040.17084544067
11	1990-10-01	2742.54697724152
12	1990-11-01	2583.81490911865
13	1990-12-01	2484.60700911865
14	1991-01-01	2447.12820000000
15	1991-02-01	2447.12820000000

Figure 1:Screenshot of the Copper Prices data

Table 2: Summary statistics

```
       V1
 Min.   : 1377
 1st Qu.: 2231
 Median : 4964
 Mean   : 4951
 3rd Qu.: 7387
 Max.   :10231
```

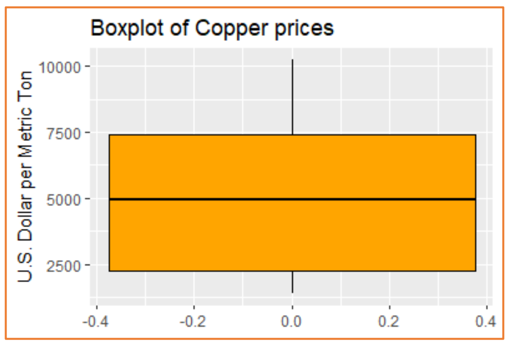


Figure 2: Boxplot giving the five number summary.

The boxplot of copper prices in figure 2 shows the median price of 4964 USD, with the interquartile range (IQR) spanning from lower value up to just 7387 USD, indicating a moderate spread in the data. However, the longer upper whisker suggests slight right skew. No visible outliers outside the whiskers which suggests that there are no extreme deviations in the data. Table 2 shows the summary statistics of the data by providing the mean, median of the data as well as the minimum and maximum value.

Research Methodology for ARMA–GARCH Model

Data Preparation

- Obtain the raw time series data to use for ARIMA.
- Transform the raw price series data to returns $r_t = \log(Y_t) - \log(Y_{t-1})$

ARIMA Model Identification and Estimation

- Stationarity testing through Box-Cox transformation and differencing.
- Identify the appropriate ARMA(p,q) process for the transformed data using ACF, PACF, and information criteria (AIC, BIC).
- Estimate ARMA parameters using maximum likelihood estimation (MLE)

Model Diagnostics

- Evaluate the adequacy of the fitted model by examining standardized residuals plot.
- Assess residual normality using Q–Q plots and tests such as Shapiro–Wilk.
- Assess correlations of residuals through the ACF graph and the Ljung-Box statistical test.

ARCH effect testing

- Assess the ACF and PACF of instantaneous non-linear transformation.
- Use the McLeod-Li statistical test.

GARCH Model Identification and Estimation

- If there are any ARCH effects, we fit the EGARCH, TGARCH, GARCH models.
- Perform EACF to identify the order of the GARCH
- Estimate GARCH parameters using maximum likelihood estimation (MLE)

Model Diagnostics

- Evaluate the adequacy of the fitted GARCH model by examining standardized residuals plot.
- Assess residual normality using Q–Q plots and tests such as Shapiro–Wilk.
- For testing correlation of returns, use the ACF along with the generalized Portmanteau test of squared and absolute residuals.

Forecasting

- Given the appropriate GARCH model, generate forecasts for the future conditional variance.

Figure 3: The Research Methodology for ARIMA-GARCH Model

Research Question

How well do ARIMA and GARCH models capture the high volatile price of copper, given the model's time varying volatility, hence we compare performances of the mean-focused ARIMA with regards to the performance of the varying-volatility-suited GARCH.

Two questions to be answered by the research

1. Given the heavy volatility of copper prices (low volatility from1990-2005 and high volatility from 2005-2025), will we ever get a fitted model that closely captures the unpredictable path of copper prices?
2. Given the past volatility, how effective are GARCH-type models in forecasting future copper price volatility given the increasing global copper demand?

ARIMA Model Identification and Estimation

Augmented Dickey-Fuller Test

H_0: The series is not stationary. H_1: The series is stationary.

```
        Augmented Dickey-Fuller Test

data:  copper_ts
Dickey-Fuller = -2.7, Lag order = 7, p-value = 0.2817
alternative hypothesis: stationary
```

Figure 3: ADF test of the Copper prices.

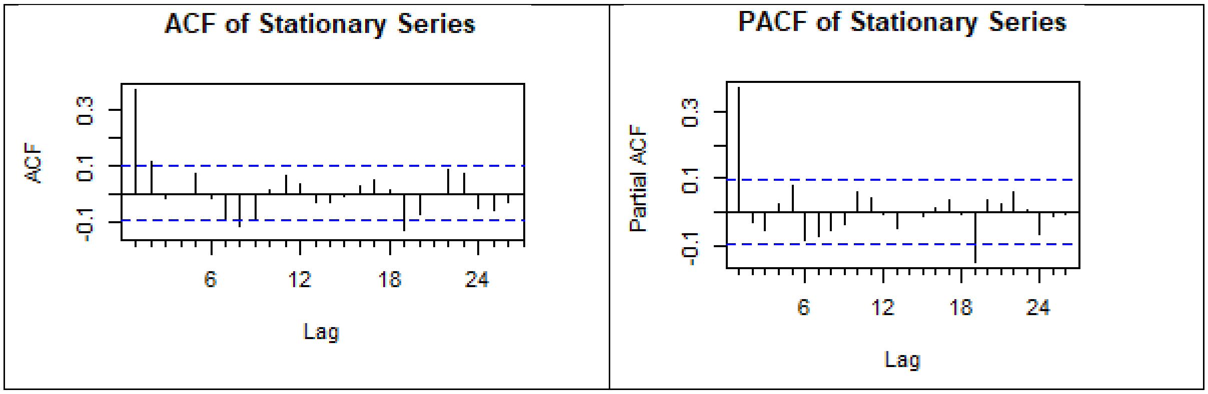
From figure 3, the ADF test of copper prices has a p-value > 0.05; thus, we fail to reject null hypothesis. Therefore, our original dataset is non-stationary.

We then conduct the Box-Cox transformation to guide us how to achieve stationarity.

Box-Cox Transformation

The estimated Box-Cox lambda is given as $\lambda = -0.2417883$. We transform the data and use the *ndiffs()* to determine order of differencing (d) to achieve stationarity. In our case the order of differencing is $d = 1$, first difference.

Table 3: ACF and PACF of Stationary Series.



The Correlogram and Partial Correlogram of our stationary data in table 3 suggest we fit a $\text{ARIMA}(2{,}1{,}1)$ model.

Augmented Dickey-Fuller Test

H_O: The series is not stationary. H_1: The series is stationary.

```
        Augmented Dickey-Fuller Test

data:  stationary
Dickey-Fuller = -7.8031, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Figure 4: ADF test of the transformed data.

Figure 4 has a p-value = 0.01 < 0.05, thus we reject null hypothesis that the series is not stationary. We achieved stationarity through data transformation; we can now fit a model.

Model Identification

Running the *auto.arima()* function in R.

```
Fitting models using approximations to speed things up...

ARIMA(2,1,2)(1,0,1)[12] with drift         : -2946.661
ARIMA(0,1,0)            with drift         : -2891.739
ARIMA(1,1,0)(1,0,0)[12] with drift         : -2954.048
ARIMA(0,1,1)(0,0,1)[12] with drift         : -2944.67
ARIMA(0,1,0)                               : -2892.49
ARIMA(1,1,0)            with drift         : -2951.239
ARIMA(1,1,0)(2,0,0)[12] with drift         : -2948.983
ARIMA(1,1,0)(1,0,1)[12] with drift         : -2952.044
ARIMA(1,1,0)(0,0,1)[12] with drift         : -2949.792
ARIMA(1,1,0)(2,0,1)[12] with drift         : -2948.761
ARIMA(0,1,0)(1,0,0)[12] with drift         : -2892.879
ARIMA(2,1,0)(1,0,0)[12] with drift         : -2951.448
ARIMA(1,1,1)(1,0,0)[12] with drift         : -2952.235
ARIMA(0,1,1)(1,0,0)[12] with drift         : -2948.908
ARIMA(2,1,1)(1,0,0)[12] with drift         : -2949.794
ARIMA(1,1,0)(1,0,0)[12]                    : -2955.532
ARIMA(1,1,0)                               : -2952.667
ARIMA(1,1,0)(2,0,0)[12]                    : -2950.28
```

```
 ARIMA(1,1,0)(1,0,1)[12]                    : -2953.556
 ARIMA(1,1,0)(0,0,1)[12]                    : -2951.274
 ARIMA(1,1,0)(2,0,1)[12]                    : -2950.095
 ARIMA(0,1,0)(1,0,0)[12]                    : -2893.725
 ARIMA(2,1,0)(1,0,0)[12]                    : -2952.884
 ARIMA(1,1,1)(1,0,0)[12]                    : -2953.71
 ARIMA(0,1,1)(1,0,0)[12]                    : -2950.183
 ARIMA(2,1,1)(1,0,0)[12]                    : -2951.253

 Now re-fitting the best model(s) without approximations...

 ARIMA(1,1,0)(1,0,0)[12]                    : -2961.709

 Best model: ARIMA(1,1,0)(1,0,0)[12]

Series: trans
ARIMA(1,1,0)(1,0,0)[12]

Coefficients:
         ar1     sar1
      0.3707   0.0363
s.e.  0.0450   0.0490

sigma^2 = 0.00005457:  log likelihood = 1483.88
AIC=-2961.77   AICc=-2961.71   BIC=-2949.61
```

Figure 5: Auto-ARIMA Overfitting and Underfitting.

Figure 5 shows the different models fitted by Auto Arima, $\text{ARIMA}(1,1,0)(1,0,0)_{12}$ being the recommended model, with an AIC of -2961.709.We then extract the AIC and BIC from the overfitting and Underfitting procedures by the Auto-ARIMA.

Table 4: AIC and BIC of different models.

	df	AIC		df	BIC
a	8	-2955.606	a	8	-2923.189
b	2	-2901.439	b	2	-2893.335
c	4	-2960.320	c	4	-2944.112
d	4	-2954.413	d	4	-2938.205
e	1	-2902.171	e	1	-2898.119
f	3	-2961.809	f	3	-2949.653
g	5	-2960.577	g	5	-2940.316
h	5	-2961.418	h	5	-2941.158
i	4	-2960.409	i	4	-2944.200
j	6	-2959.876	j	6	-2935.564
k	3	-2899.997	k	3	-2887.840
l	5	-2958.678	l	5	-2938.417
m	5	-2958.601	m	5	-2938.341
n	4	-2954.340	n	4	-2938.131
o	6	-2956.398	o	6	-2932.086
p	3	-2961.766	p	3	-2949.610
q	2	-2963.216	q	2	-2955.112
r	4	-2961.938	r	4	-2945.730
s	4	-2962.841	s	4	-2946.633
t	3	-2961.859	t	3	-2949.703
u	5	-2961.266	u	5	-2941.005
v	2	-2900.820	v	2	-2892.716
w	4	-2960.097	w	4	-2943.888
x	4	-2960.028	x	4	-2943.820
y	3	-2955.600	y	3	-2943.444
z	5	-2957.845	z	5	-2937.585

From table 4 we see that model Q is the best model with the lowest AIC and BIC; thus, we will then fit model $\mathrm{ARIMA}(1,1,0)$ without a drift.

Parameter Estimation

$\mathrm{ARIMA}(1,1,0)$:

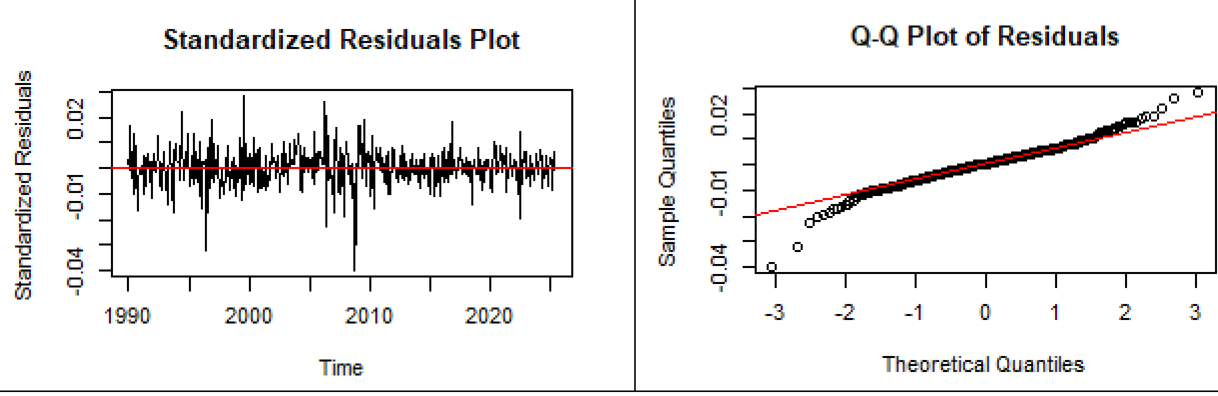
```
Coefficients:
         ar1
      0.3709
s.e.  0.0450

sigma^2 = 0.00005452:  log likelihood = 1483.61
AIC=-2963.22   AICc=-2963.19   BIC=-2955.11
```

Figure 6: Parameter estimates of the chosen model.

$Y_t^{0.24} = 0.3709Y_{t-1}^{0.24} + e_t$ with noise error of 0.0000542, is our fitted model. Parameter estimates are significant.

Residual Analysis

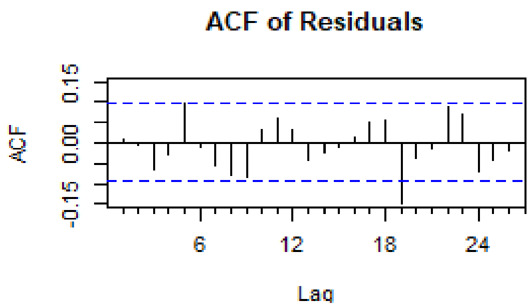
Table 5: Standardized residuals Plot, QQ-Plot and Shapiro-Wilk test.




```
Shapiro-Wilk normality test

data:  resids
W = 0.96586, p-value = 0.00000002107
```

From table 5 above the standardized residuals do not show any apparent pattern but are fairly "random" and all lie within ± 3 interval except for a reasonably large residual in 2008. Most of the points on the QQ plot lie on the line expect largely for the lower tail, but then residuals may still have a normal distribution. It is then observed that the Shapiro-Wilk test in table 5 indicates that we reject null hypothesis that the residuals are normally distributed, since p-value < 0.05.

Table 6: ACF and Ljung-Box test for autocorrelation of residuals.




```
	Box-Ljung test

data:  resids
X-squared = 29.499, df = 20, p-value = 0.07839
```

From table 6, we observe that the only significant residual is at lag 19 but the rest are insignificant. The Ljung-Box test on the right of table 6 suggests we do not reject the null hypothesis that residuals are independent since p-value is 0.07839 > 0.05.

Based on the above, it follows that the fitted model may not be appropriate, particularly due to non-normal residuals.

ARCH effect testing

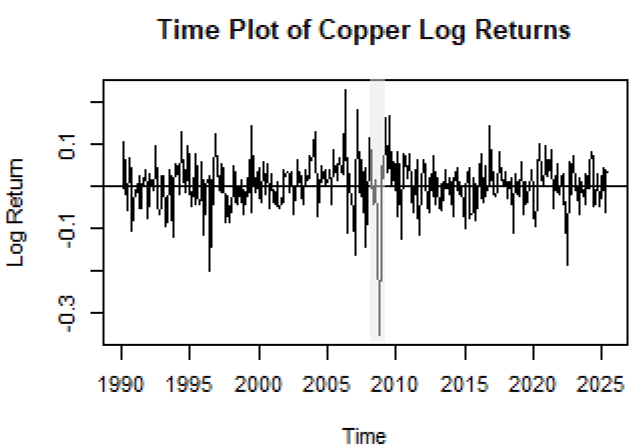
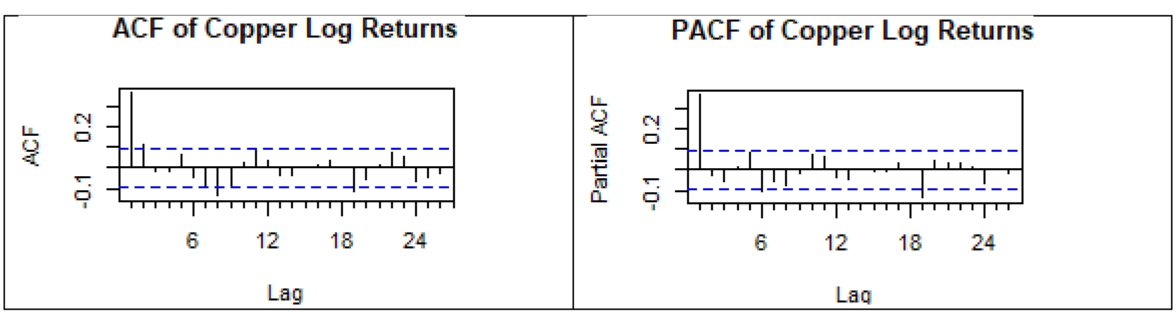


Figure 7: Plot of log returns of copper prices.

Figure 7 is a plot of log returns of copper prices, the returns depict alternating quiet and volatile periods, volatility clustering. The shaded part being due to the 2008 global financial crises.

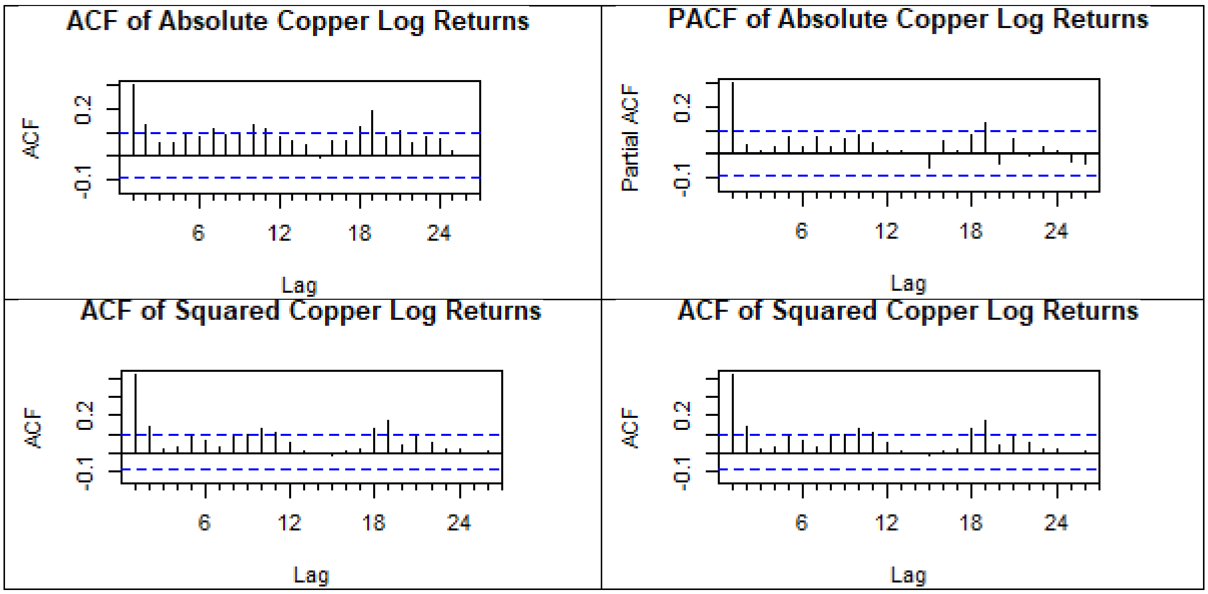
Table 7: ACF and PACF of the returns.



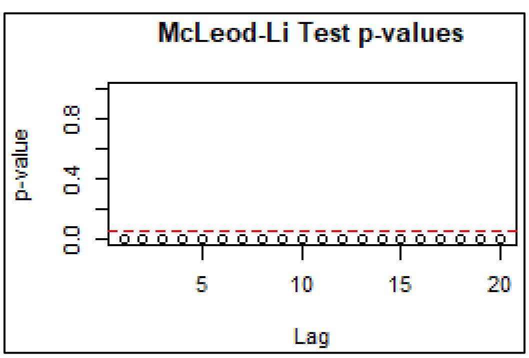
The ACF and PACF in table 7 suggest the returns r_t are significant on lag 1 and 2 for the ACF, and significant at lag 1 for the PACF, hence there is autocorrelation in the returns. Thus, the ACF and PACF do suggest that the returns are not i.i.d., due to volatility clustering.

Let us confirm that the returns are truly dependent by doing a nonlinear instantaneous transformation such as taking the absolute values and squaring the returns, which tend to preserve independence.

Table 8: ACF and PACF for the nonlinear instantaneous transformation results.



From table 8 we see that after plotting the ACF and PACF of the absolute and squared values of the returns, many more lags have become significant, thus reinforcing the suggestion that the returns are autocorrelated, implying possible ARCH in the returns.

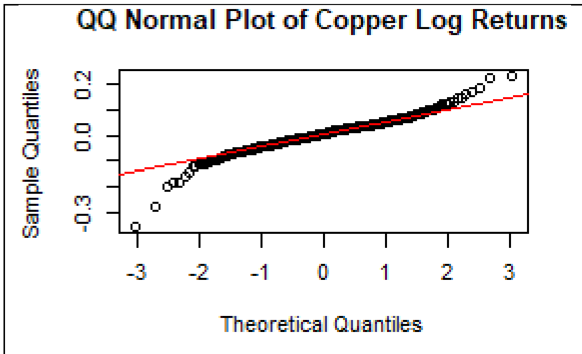


The Mcleod-Li test, in figure 8, suggests we reject the null hypothesis; H_0: There is no ARCH. We then conclude that the returns have conditional heteroscedasticity.

Figure 8: McLeod-Li test for autocorrelation.

We will next look at the distribution pattern of the returns so we can establish whether a GARCH model is appropriate to fit.

Table 9: QQ plot and Shapiro-Wilk test for distribution of returns.




```
        Shapiro-Wilk normality test

data:  rt
W = 0.94557, p-value = 0.0000000002216
```

The QQ plot in table 9 suggests to us that the returns may have a thicker upper tail and lighter left tail than that of a normal distribution, thus may be skewed to the right. The Shapiro-Wilk test has p-value < 0.05, hence we then reject the null hypothesis that returns are normally distributed. The sample skewness of the copper returns is equal to -0.6219191 and the sample kurtosis in excess of 3 is equal to 4.83966.

We then conclude that the copper returns are dependent , and have volatility clustering, and are non-normally distributed, highly leptokurtic with uneven tails. We can then model a GARCH model to capture the properties of our data.

Fitting different models from the GARCH family

We fit models manually using the *optim()* function in R.

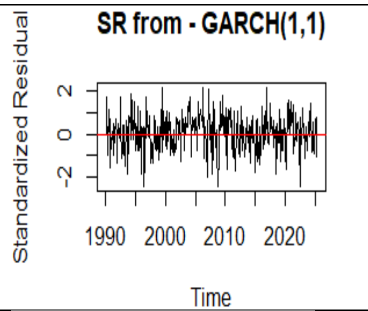
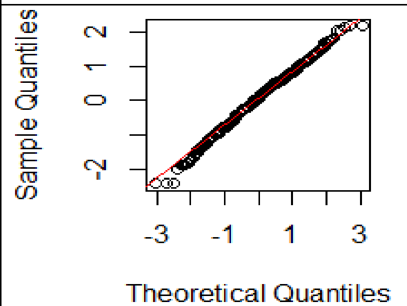
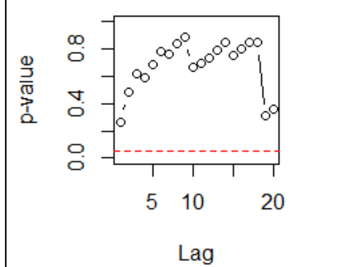
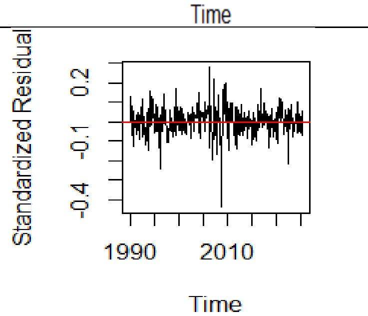
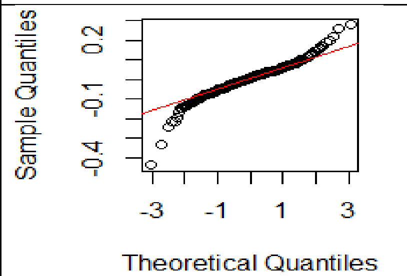
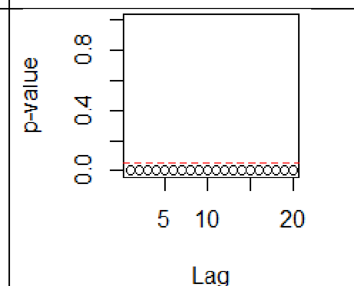
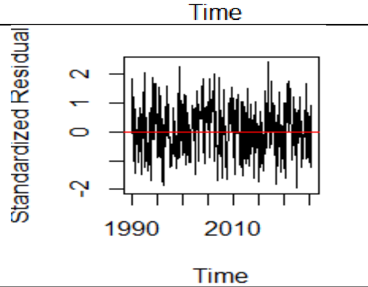
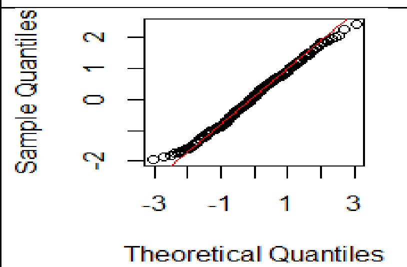
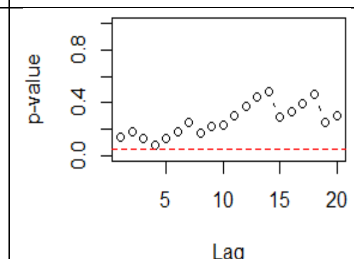
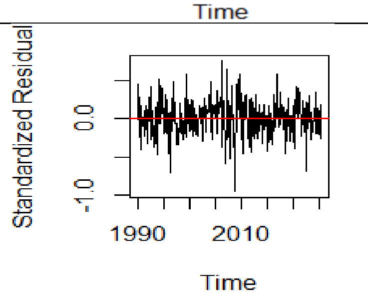
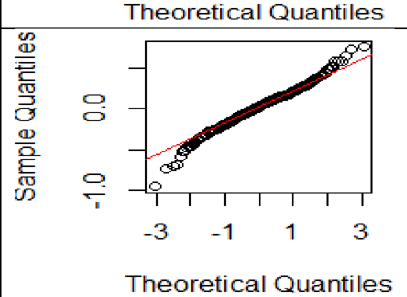
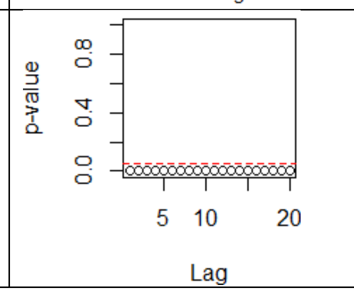
Table 10: Parameter Estimates for the families of GARCH.

Model	**Parameter**	**Estimate**	**Std. Error**
GARCH(1,1)	ω	0.00019232	0.00007609
	α	0.08741933	0.02685077
	β	0.85631838	0.04142544
EGARCH(1,1)	ω	-0.3585834	0.4454619
	α	0.2277245	0.1191142
	γ	-0.00394489	0.04523508
	β	0.9361736	0.0781384
GJR-GARCH(1,1)	ω	0.0001410219	0.000002776706
	α	0.1482796	0.004495857
	γ	0.8237968	0.007260513
	β	-0.006094916	0.002341653
APARCH(1,1)	ω	0.00173009	0.00002856728
	α	0.09997467	0.02893322
	γ	0.7994487	0.04432005
	β	0.09999985	0.1432076
	δ	1.499786	0.09804682

From table 10 GARCH(1,1) and GJR-GARCH(1,1) models have significant estimated parameters, whereas EGARCH(1,1) and APARCH(1,1) models have insignificant parameters being γ and β respectively. γ is only negative for EGARCH(1,1), which is typical for a gamma in financial data, since negative returns increase volatility more than positive returns.

Model Diagnostic

Table 11: Residual analysis of the fitted models.

Model: AIC|BIC|P-Value	**Standardized residuals (SR)**	**QQ Plot**	**Portmanteau test of SR2**
<u>GARCH(1,1)</u> AIC: -1233.445 BIC: -1221.289 Shapiro-Wilk: p-value = 0.2251			
<u>EGARCH(1,1)</u> AIC: -1237.274 BIC: -1225.117 Shapiro-Wilk: p-value = 8.125×10^{-12}			
<u>GJR-GARCH(1,1)</u> AIC: -1235.155 BIC: -1218.947 Shapiro-Wilk: p-value = 0.04201			
<u>APARCH(1,1)</u> AIC: -1234.349 BIC: -1214.088 Shapiro-Wilk: p-value = 4.189×10^{-4}			

From table 11, the standardized residuals plot of all the models seems to be random, with no apparent pattern and outliers (± 3 CI). Form the table the QQ plot of GARCH(1,1) and GJR-GARCH(1,1) have most of their points on the line; whereas EGARCH(1,1) and APARCH(1,1) show points that deviate from the line on the tails, hence squared residuals might not be normal, Shapiro-Wilk test indicates GARCH(1,1) is the only model with normality (p-value > 0.05), as we fail to reject null hypothesis that residuals are normally distributed. The portmanteau test of the squared residuals of the APARCH (1,1) and EGARCH(1,1) have p-values lower than 0.05; hence we reject the null hypothesis that the residuals are uncorrelated (independent).

Based on the Analysis GARCH(1,1) and GJR-GARCH(1,1) seem to be the best models; but EGARCH(1,1) has the smallest AIC of -1237.274. despite it having non-normal and correlated squared residuals.

According to Grok AI, EGARCH(1,1)'s high and significant β persistence parameter better captures volatility persistence, which dominates copper price behaviour, besides having a weak asymmetry due to the insignificant γ asymmetry parameter, but γ is negative for EGARCH model, which is typical for financial data, its significance can be improve perhaps increasing data size.

Conditional Variance Forecasting

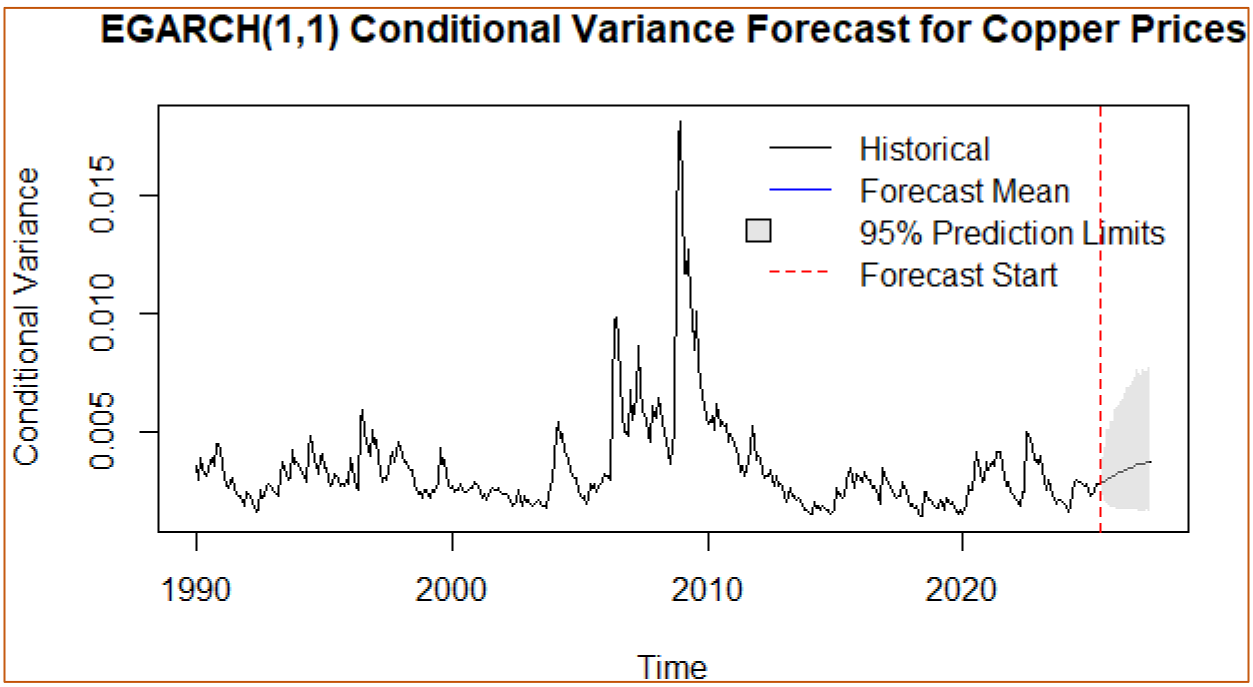


Figure 9: Conditional variance plot of copper prices with a 24-month forecast.

Figure 9 is a conditional variance plot that has 24-month conditional variance forecast, with 955 prediction limits which are slightly narrow at the forecast start period and greatly wider due to the extreme uncertainty behind the highly sensitive copper prices. The forecasted conditional variance approaches the long-run variance of $\sigma^2 = \exp\left(\frac{\omega}{1-\beta}\right) = \exp(-5.6186) = 0.003631.$

In closing remarks, we were able to fit a GARCH model, particularly an EGARCH(1,1) despite having non-normal and correlated residuals squared. Due to its desirable characteristics such as high ability of model persistence and a negative asymmetry parameter despite it being insignificant.