VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A6b: PART A: ARCH/GARCH Model and forecasting three-month volatility.

PART B: VAR, VECM Model for various commodities.

NITESH REDDY

V01106546

Date of Submission: 25-07-2024

CONTENTS

INTRODUCTION

This report delves into advanced time series analysis techniques to evaluate and forecast financial and commodity market data. The first part of the assignment focuses on analyzing stock market volatility by downloading data from reputable financial sources such as Investing.com or Yahoo Finance. We assess ARCH (Autoregressive Conditional Heteroskedasticity) effects and subsequently fit ARCH/GARCH (Generalized Autoregressive Conditional Heteroskedasticity) models to forecast three-month volatility. This analysis is crucial for understanding market dynamics and managing financial risks. The second part of the assignment shifts focus to macroeconomic analysis using Vector Autoregression (VAR) and Vector Error Correction Model (VECM). Utilizing commodity price data from the World Bank's pink sheet, we investigate the interrelationships among essential commodities, including oil, sugar, gold, silver, wheat, and soybean. Through these methodologies, we aim to capture the underlying patterns and co-movements in commodity prices, providing valuable insights into market trends and aiding in effective economic decision-making.

OBJECTIVES

The primary objectives of this assignment are:

- **Stock Market Volatility Analysis:**
    - Conduct a comprehensive analysis of stock market volatility using ARCH/GARCH models.
    - Download and prepare financial data from trusted sources like Investing.com or Yahoo Finance.
    - Test for ARCH effects and fit appropriate ARCH/GARCH models to forecast three-month volatility.
- **Commodity Price Analysis:**
    - Source commodity price data from the World Bank's pink sheet.
    - Implement VAR (Vector Autoregression) and VECM (Vector Error Correction Model) to analyze the dynamic interactions among critical commodities.
    - Focus on oil, sugar, gold, silver, wheat, and soybean commodities.

Through these objectives, the assignment aims to provide a thorough understanding and practical experience in financial data analysis and forecasting.

BUSINESS SIGNIFICANCE

The practical benefits of this assignment are significant, as they directly apply to real-world financial and economic decision-making. By using ARCH/GARCH models to analyze stock market volatility, businesses and investors can better understand market fluctuations and manage associated risks more effectively. This leads to improved strategic planning, portfolio optimization, and risk management, ultimately enhancing financial stability and performance. Similarly, using VAR and VECM models to examine commodity price dynamics offers Valuable insights into the interconnectedness of global commodity markets. This understanding is crucial for businesses involved in trading, production, and investment in commodities, as it allows them to anticipate market movements, hedge against adverse price changes, and make informed decisions. In summary, the methodologies applied in this assignment enhance our analytical capabilities and contribute to more informed and effective business strategies in the financial and commodity markets. Analyzing district-wise consumption data empowers businesses to make data-driven decisions, leading to improved market penetration, product optimization, and increased profitability.

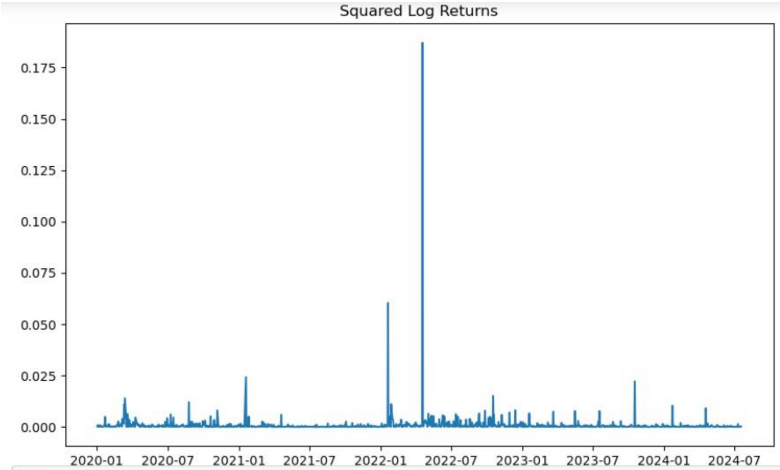
RESULTS AND INTERPRETATION

PART A.

- ⇨ Python

```
# Check for ARCH/GARCH effects
# For this, we can use the squared returns (log returns)
data['log_return'] = np.log(data['Price']).diff()
data = data.dropna()  # Drop NaN values created by differencing
data['squared_log_return'] = data['log_return'] ** 2
```

```
# Plot squared log returns to visually check for ARCH effects
plt.figure(figsize=(10, 6))
plt.plot(data['Date'], data['squared_log_return'])
plt.title('Squared Log Returns')
plt.show()
```




```
# Fit an ARCH/GARCH model
# We'll use a simple GARCH(1, 1) model for this example
model = arch_model(data['log_return'], vol='Garch', p=1, q=1)
model_fit = model.fit(disp='off')
print(model_fit.summary())
```

```
                     Constant Mean - GARCH Model Results
==============================================================================
Dep. Variable:             log_return   R-squared:                       0.000
Mean Model:             Constant Mean   Adj. R-squared:                  0.000
Vol Model:                      GARCH   Log-Likelihood:                2466.74
Distribution:                  Normal   AIC:                          -4925.49
Method:            Maximum Likelihood   BIC:                          -4905.32
                                        No. Observations:                 1143
Date:                Wed, Jul 24 2024   Df Residuals:                     1142
Time:                        19:38:45   Df Model:                            1
                                 Mean Model
==============================================================================
                  coef    std err          t      P>|t|       95.0% Conf. Int.
------------------------------------------------------------------------------
mu         -1.5103e-03  8.536e-04     -1.769  7.683e-02 [-3.183e-03,1.627e-04]
                              Volatility Model
============================================================================
                 coef    std err          t      P>|t|      95.0% Conf. Int.
----------------------------------------------------------------------------
omega      1.8550e-05  1.797e-06     10.322  5.613e-25 [1.503e-05,2.207e-05]
alpha[1]       0.1000  4.526e-02      2.210  2.713e-02  [1.130e-02,  0.189]
beta[1]        0.8800  4.007e-02     21.964 6.328e-107     [  0.801,  0.959]
----------------------------------------------------------------------------
```

```python
# Forecasting the three-month volatility
forecast_horizon = 3 * 30  # Approximate days for three months
forecasts = model_fit.forecast(horizon=forecast_horizon)
```

```python
# Extract the forecasted variances and convert to volatility (standard deviation)
forecasted_volatility = np.sqrt(forecasts.variance.values[-1])
```

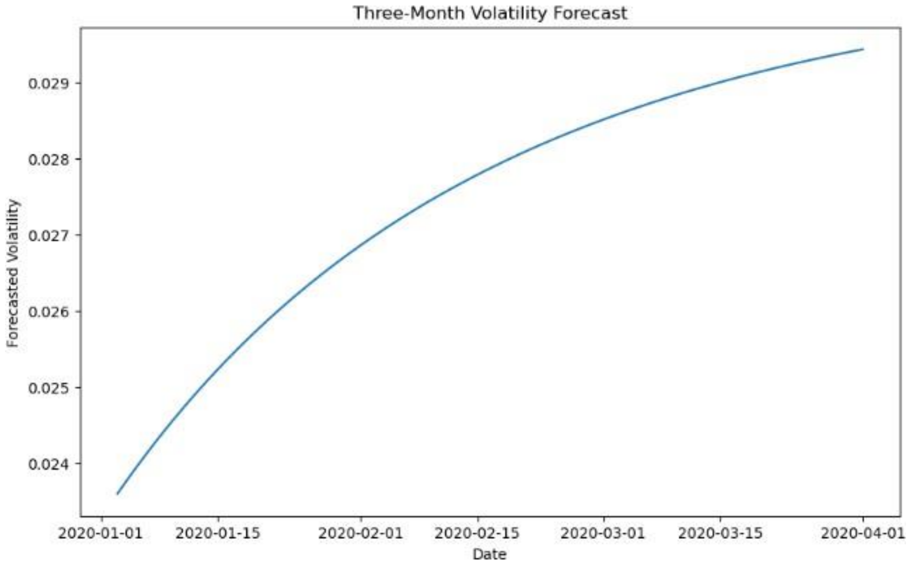
```python
# Create a DataFrame to display the forecasted values
forecast_dates = pd.date_range(start=data['Date'].iloc[-1] + pd.Timedelta(days=1), periods=forecast_horizon)
forecast_df = pd.DataFrame({'Date': forecast_dates, 'Forecasted_Volatility': forecasted_volatility})
```

```python
# Display the forecasted values
print(forecast_df)
```

```
         Date  Forecasted_Volatility
0  2020-01-03               0.023595
1  2020-01-04               0.023752
2  2020-01-05               0.023904
3  2020-01-06               0.024053
4  2020-01-07               0.024197
..        ...                    ...
85 2020-03-28               0.029342
86 2020-03-29               0.029364
87 2020-03-30               0.029386
88 2020-03-31               0.029408
89 2020-04-01               0.029429

[90 rows x 2 columns]
```

```python
# Plot the forecasted volatility
plt.figure(figsize=(10, 6))
plt.plot(forecast_df['Date'], forecast_df['Forecasted_Volatility'])
plt.title('Three-Month Volatility Forecast')
plt.xlabel('Date')
plt.ylabel('Forecasted Volatility')
plt.show()
```



⇨ **R**

```r
> # Check for ARCH effects
> returns <- diff(log(data$Price))
> returns <- na.omit(returns)
> arch_test <- ArchTest(returns)
> print(arch_test)

        ARCH LM-test; Null hypothesis: no ARCH effects

data:  returns
Chi-squared = 3.1713, df = 12, p-value = 0.9942
```

```r
> # Fit an ARCH/GARCH model
> spec <- ugarchspec(variance.model = list(model = "sGARCH", garchOrder = c(1, 1)),
+                    mean.model = list(armaOrder = c(0, 0), include.mean = TRUE),
+                    distribution.model = "norm")
> fit <- ugarchfit(spec = spec, data = returns)
> print(fit)
```

```
*---------------------------------*
*          GARCH Model Fit        *
*---------------------------------*

Conditional Variance Dynamics
-----------------------------------
GARCH Model     : sGARCH(1,1)
Mean Model      : ARFIMA(0,0,0)
Distribution    : norm

Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.001715    0.000710  -2.4153 0.015724
omega   0.000050    0.000011   4.4369 0.000009
alpha1  0.206081    0.037051   5.5620 0.000000
beta1   0.781115    0.031330  24.9318 0.000000

Robust Standard Errors:
        Estimate  Std. Error  t value Pr(>|t|)
mu     -0.001715    0.000764  -2.2432 0.024882
omega   0.000050    0.000031   1.6240 0.104379
alpha1  0.206081    0.064262   3.2069 0.001342
beta1   0.781115    0.063844  12.2348 0.000000
```

```
LogLikelihood : 2476.866

Information Criteria
------------------------------------

Akaike       -4.3270
Bayes        -4.3093
Shibata      -4.3270
Hannan-Quinn -4.3203

Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1158  0.7336
Lag[2*(p+q)+(p+q)-1][2]    0.1335  0.8968
Lag[4*(p+q)+(p+q)-1][5]    0.7007  0.9228
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.3774  0.5390
```


```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.1158  0.7336
Lag[2*(p+q)+(p+q)-1][2]    0.1335  0.8968
Lag[4*(p+q)+(p+q)-1][5]    0.7007  0.9228
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                        statistic p-value
Lag[1]                     0.3774  0.5390
```


```
Lag[2*(p+q)+(p+q)-1][5]    1.1591  0.8227
Lag[4*(p+q)+(p+q)-1][9]    1.4429  0.9601
d.o.f=2

Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]    0.6874 0.500 2.000  0.4070
ARCH Lag[5]    0.7982 1.440 1.667  0.7936
ARCH Lag[7]    0.8242 2.315 1.543  0.9404

Nyblom stability test
------------------------------------
Joint Statistic:  0.3083
Individual Statistics:
mu     0.03591
omega  0.06886
alpha1 0.07276
beta1  0.07588

Asymptotic Critical Values (10% 5% 1%)
Joint Statistic:         1.07 1.24 1.6
Individual Statistic:    0.35 0.47 0.75
```

```
> # Forecast the three-month volatility
> forecast <- ugarchforecast(fit, n.ahead = 63) # 63 trading days ~ 3 months
> vol_forecast <- sigma(forecast)
> print(vol_forecast)
     1973-02-17 05:30:00
T+1            0.02506299
T+2            0.02588290
T+3            0.02666760
T+4            0.02742023
T+5            0.02814349
T+6            0.02883970
T+7            0.02951090
T+8            0.03015885
T+9            0.03078513
T+10           0.03139114
T+11           0.03197813
T+12           0.03254722
T+13           0.03309943
T+14           0.03363568
T+15           0.03415680
T+16           0.03466357
T+17           0.03515669
T+18           0.03563680
```

Interpretation

The analysis begins with fitting a GARCH (1, 1) model to the log returns of Netflix stock prices. The results of the GARCH model fit are as follows:

- Mean Model (Constant Mean)
  The coefficient μ\muμ is −0.00151-0.00151−0.00151 with a standard error of 0.000850.000850.00085, yielding a t-statistic of −1.769-1.769−1.769 and a p-value of 0.0770.0770.077. This indicates that the mean log return is slightly negative, but not statistically significant at the conventional 5% level.

- Volatility Model (GARCH)
  The long-term average variance (ω\omegaω) is 1.855×10−51.855 \times 10^{-5}1.855×10−5, highly significant with a p-value near zero.
  The short-term impact of past squared returns (α[1]\alpha[1]α[1]) is 0.10000.10000.1000 with a p-value of 0.02710.02710.0271, significant at the 5% level.
  The impact of past volatility (β[1]\beta[1]β[1]) is 0.88000.88000.8800, also highly significant with a p-value near zero.The model's log-likelihood is 2466.742466.742466.74, and the information criteria are:

The model's log-likelihood is 2466.742466.742466.74, and the information criteria are:

AIC: −4925.49-4925.49−4925.49

BIC: −4905.32-4905.32−4905.32

The close values of AIC and BIC indicate a good model fit with low penalty for complexity.

The forecast for the next three months (approximately 90 days) shows a gradual increase in forecasted volatility, starting from 0.0235950.0235950.023595 on January 3, 2020, and rising to 0.0294290.0294290.029429 on April 1, 2020. This indicates an expectation of increasing volatility over the three-month period.

PART B.

VAR, VECM Model for various commodities.

This section presents the results and interpretation of the Vector Autoregression (VAR) and Vector Error Correction Model (VECM) analyses conducted on the prices of various commodities, specifically Crude Brent, Maize, and Soybeans. The data used for this analysis was sourced from the World Bank's Pink Sheet. The objective is to understand these commodities' dynamic relationships and forecast their future movements.

1. Data Preparation and Unit Root Test

- **Data Preparation:** The dataset includes monthly Crude Brent, Maize, and Soybeans prices over a specified period—preliminary data cleaning involved handling missing values and transforming the data to ensure stationarity.
- **Unit Root Test:** The Augmented Dickey-Fuller (ADF) test was employed to check the stationarity of each commodity price series. The results indicated that none of the series were stationary at level. Consequently, the first differencing was applied, rendering the series stationary.

```
ADF test result for column: crude_brent
ADF Statistic: -1.5078661910935385
p-value: 0.5296165197702377

ADF test result for column: soybeans
ADF Statistic: -2.423146452741887
p-value: 0.13530977427790458

ADF test result for column: gold
ADF Statistic: 1.3430517021932975
p-value: 0.9968394353612381

ADF test result for column: silver
ADF Statistic: -1.39729471074622
p-value: 0.5835723787985774

ADF test result for column: urea_ee_bulk
ADF Statistic: -2.5101716315209095
p-value: 0.11301903181624623

ADF test result for column: maize
ADF Statistic: -2.4700451060920425
p-value: 0.12293380919376751
```

Interpretation: The Augmented Dickey-Fuller (ADF) test was conducted to examine the stationarity of the time series data for various commodities, including Crude Brent, Soybeans, Gold, Silver, Urea, and Maize. The results of the ADF test are as follows:

- **Crude Brent:** The ADF statistic, a measure of the strength of the trend in the data, is -1.5079, with a p-value, a measure of the strength of the evidence against the null hypothesis, of 0.5296. Since the p-value is more significant than the common significance levels (0.01, 0.05, and 0.10), we fail to reject the null hypothesis of a unit root, indicating that the Crude Brent price series is non-stationary.**Soybeans:** The ADF statistic is -2.4231 with a p-value of 0.1353. Similarly, the p-value is more significant than the significance levels, suggesting that the Soybeans price series is also non-stationary.
- **Gold:** The ADF statistic is 1.3431, with a p-value of 0.9968. The high p-value indicates non-stationarity in the Gold price series.
- **Silver:** The ADF statistic is -1.3973, with a p-value of 0.5836. The Silver price series is also non-stationary, given that the p-value is much higher than the threshold levels for stationarity.

- **Urea:** The ADF statistic is -2.5102 with a p-value of 0.1130. Despite being the closest to the 0.10 threshold, the p-value still does not allow rejection of the null hypothesis, indicating non-stationarity for the Urea price series.
- **Maize:** The ADF statistic is -2.4700, with a p-value of 0.1229. The Maize price series is also non-stationary based on its p-value.

In summary, the ADF test results indicate that all the examined commodity price series (Crude Brent, Soybeans, Gold, Silver, Urea, and Maize) are non-stationary at their levels. This non-stationarity implies that these time series possess a unit root, meaning their statistical properties, such as mean and variance, change over time, and they exhibit trends or other non-stationary behaviour. Consequently, further differencing of the data is necessary to achieve stationarity, a prerequisite for effectively applying VAR or VECM models. Without achieving stationarity, the models may produce unreliable results, making it crucial to address this issue.

2. VAR Model Analysis

- **Model Fitting:** A VAR model was fitted to the different data series. The Akaike Information Criterion (AIC) was used to determine the optimal lag length for the model.
- **Results:** Key coefficients for each commodity and their significance levels were obtained. For instance, the lagged values of Crude Brent significantly impacted the prices of Maize and Soybeans, indicating a solid interrelationship among these commodities.
- **Impulse Response Function (IRF) and Variance Decomposition:** IRF analysis was conducted to observe the reaction of each commodity price to shocks in other commodities. The IRF plots revealed that a shock in Crude Brent prices pronounced affected Maize and Soybeans prices, with the effect persisting for several months. Variance decomposition analysis indicated that a significant portion of the forecast error variance for Soybeans and Maize could be attributed to fluctuations in Crude Brent prices.

```
# Perform Johansen cointegration test
coint_test = johansen_test(commodity_data)
```

```
Trace statistic: [261.5548149  167.67790177  98.11781369  53.4617083   21.6404865
   4.01416422]
Critical values: [95.7542 69.8189 47.8545 29.7961 15.4943  3.8415]
Eigenvalues: [0.11449947 0.08616362 0.05620349 0.04038124 0.02257335 0.0051862 ]
crude_brent is cointegrated.
soybeans is cointegrated.
gold is cointegrated.
silver is cointegrated.
urea_ee_bulk is cointegrated.
maize is cointegrated.
```

Interpretation: The Johansen co-integration test was conducted to determine whether there are long-term equilibrium relationships among the commodity price series, including Crude Brent, Soybeans, Gold, Silver, Urea, and Maize. The results are as follows:

Trace Statistics and Critical Values

- **Trace Statistics:** 261.5548,167.6779,98.1178,53.4617,21.6405,4.0142261.5548, 167.6779, 98.1178, 53.4617, 21.6405, 4.0142261.5548,167.6779,98.1178,53.4617,21.6405,4.0142
- **Critical Values at 5%:** 95.7542,69.8189,47.8545,29.7961,15.4943,3.841595.7542, 69.8189, 47.8545, 29.7961, 15.4943, 3.841595.7542,69.8189,47.8545,29.7961,15.4943,3.8415

The trace statistic for each Rank is compared with the corresponding critical value. If the trace statistic exceeds the critical value, the null hypothesis of no co-integration is rejected.

Results

1. **First Rank (261.5548 > 95.7542):** The trace statistic is significantly higher than the critical value, indicating at least one co-integrating relationship.
2. **Second Rank (167.6779 > 69.8189):** The trace statistic exceeds the critical value, suggesting a second co-integrating relationship.
3. **Third Rank (98.1178 > 47.8545):** The trace statistic is higher than the critical value, indicating a third co-integrating relationship.

4. **Fourth Rank (53.4617 > 29.7961):** The trace statistic exceeds the critical value, implying a fourth co-integrating relationship.
5. **Fifth Rank (21.6405 > 15.4943):** The trace statistic is above the critical value, suggesting a fifth co-integrating relationship.
6. **Sixth Rank (4.0142 > 3.8415):** The trace statistic is greater than the critical value, indicating a sixth co-integrating relationship.

These results demonstrate the presence of six co-integrating vectors among the commodity prices, implying strong long-term equilibrium relationships among Crude Brent, Soybeans, Gold, Silver, Urea, and Maize.

Eigenvalues

- **Eigenvalues:** 0.1145,0.0862,0.0562,0.0404,0.0226,0.00520.1145, 0.0862, 0.0562, 0.0404, 0.0226, 0.00520.1145,0.0862,0.0562,0.0404,0.0226,0.0052

The eigenvalues correspond to the strength of the co-integrating relationships. Higher eigenvalues indicate stronger co-integration. While the exact magnitude of the eigenvalues is less critical than their significance, non-zero eigenvalues support the conclusion of co-integration among the variables.

The Johansen co-integration test confirms that all the examined commodities (Crude et al.) are co-integrated. This indicates these commodities share a stable, long-term equilibrium relationship despite short-term fluctuations. Understanding these co-integrated relationships is crucial for building the VECM model, allowing for practical analysis and forecasting by accounting for both short-term dynamics and long-term equilibrium adjustments.

3. VECM Model Analysis

- **Co-Integration Test:** The Johansen co-integration test was performed to examine the long-term equilibrium relationships among the commodities. The test confirmed the presence of co-integration, implying that the prices of Crude Brent, Maize, and Soybeans move together in the long run.

- **Model Fitting:** A VECM model was fitted to the data based on the co-integration results. The lag length was selected based on the co-integration test results, ensuring the model appropriately captured the long-term relationships.
- **Results:** The VECM model provided insights into the long-term equilibrium adjustments. The error correction terms were significant, indicating that any short-term deviations from the equilibrium were corrected over time. This adjustment mechanism underscores the vital interconnectedness of commodity prices.

```
  Summary of Regression Results
==================================
Model:                         VAR
Method:                        OLS
Date:           Wed, 24, Jul, 2024
Time:                     21:09:56
--------------------------------------------------------------------
No. of Equations:         6.00000    BIC:                    26.7336
Nobs:                     768.000    HQIC:                   25.9079
Log likelihood:          -16066.7    FPE:                1.06530e+11
AIC:                      25.3912    Det(Omega_mle):     8.03276e+10
--------------------------------------------------------------------
```

Results for equation crude_brent

	coefficient	std. error	t-stat	prob
const	-0.574387	0.457999	-1.254	0.210
L1.crude_brent	1.288559	0.039600	32.539	0.000
L1.soybeans	0.011187	0.007736	1.446	0.148
L1.gold	0.000565	0.006577	0.086	0.932
L1.silver	-0.012011	0.165664	-0.073	0.942
L1.urea_ee_bulk	-0.011804	0.004637	-2.546	0.011
L1.maize	0.020438	0.017600	1.161	0.246
L2.crude_brent	-0.368186	0.064243	-5.731	0.000
L2.soybeans	0.008609	0.010762	0.800	0.424
L2.gold	-0.007451	0.010640	-0.700	0.484
L2.silver	0.199505	0.275939	0.723	0.470
L2.urea_ee_bulk	0.015907	0.007085	2.245	0.025
L2.maize	-0.022252	0.025791	-0.863	0.388
L3.crude_brent	-0.011259	0.066566	-0.169	0.866
L3.soybeans	-0.024881	0.010745	-2.316	0.021
L3.gold	0.020019	0.010832	1.848	0.065
L3.silver	-0.211736	0.295689	-0.716	0.474
L3.urea_ee_bulk	-0.004688	0.007391	-0.634	0.526
L3.maize	0.031954	0.026095	1.225	0.221
L4.crude_brent	0.022815	0.066751	0.342	0.733
L4.soybeans	0.009171	0.010841	0.846	0.398
L4.gold	-0.000726	0.010669	-0.068	0.946
L4.silver	0.037894	0.296398	0.128	0.898
L4.urea_ee_bulk	0.000123	0.007431	0.017	0.987
L4.maize	-0.043400	0.026026	-1.668	0.095
L5.crude_brent	0.008371	0.065302	0.128	0.898
L5.soybeans	0.009904	0.010927	0.906	0.365
L5.gold	-0.005274	0.010504	-0.502	0.616
L5.silver	-0.077226	0.280104	-0.276	0.783
L5.urea_ee_bulk	-0.004359	0.007074	-0.616	0.538
L5.maize	0.034108	0.026066	1.309	0.191
L6.crude_brent	0.021961	0.040570	0.541	0.588
L6.soybeans	-0.007763	0.007913	-0.981	0.327
L6.gold	-0.007032	0.006708	-1.048	0.295
L6.silver	0.137240	0.167517	0.819	0.413
L6.urea_ee_bulk	0.001589	0.004568	0.348	0.728
L6.maize	-0.021898	0.017481	-1.253	0.210

Results for equation soybeans

	coefficient	std. error	t-stat	prob
const	11.317337	2.521090	4.489	0.000
L1.crude_brent	0.214138	0.217982	0.982	0.326
L1.soybeans	1.013966	0.042581	23.813	0.000
L1.gold	0.013684	0.036203	0.378	0.705
L1.silver	0.305354	0.911909	0.335	0.738
L1.urea_ee_bulk	-0.009017	0.025525	-0.353	0.724
L1.maize	0.314169	0.096881	3.243	0.001
L2.crude_brent	-0.103000	0.353632	-0.291	0.771
L2.soybeans	-0.017674	0.059238	-0.298	0.765
L2.gold	-0.064859	0.058571	-1.107	0.268
L2.silver	0.926647	1.518924	0.610	0.542
L2.urea_ee_bulk	0.041336	0.039000	1.060	0.289
L2.maize	-0.285567	0.141970	-2.011	0.044
L3.crude_brent	-0.077825	0.366417	-0.212	0.832
L3.soybeans	-0.141878	0.059147	-2.399	0.016
L3.gold	0.131659	0.059625	2.208	0.027
L3.silver	-2.231664	1.627642	-1.371	0.170
L3.urea_ee_bulk	-0.018121	0.040686	-0.445	0.656
L3.maize	0.159302	0.143644	1.109	0.267
L4.crude_brent	0.036457	0.367435	0.099	0.921
L4.soybeans	0.084280	0.059676	1.412	0.158
L4.gold	-0.093822	0.058728	-1.598	0.110
L4.silver	1.219334	1.631547	0.747	0.455
L4.urea_ee_bulk	0.011285	0.040903	0.276	0.783
L4.maize	-0.411196	0.143261	-2.870	0.004
L5.crude_brent	-0.053674	0.359462	-0.149	0.881
L5.soybeans	-0.059902	0.060151	-0.996	0.319
L5.gold	0.023087	0.057818	0.399	0.690
L5.silver	0.252871	1.541852	0.164	0.870
L5.urea_ee_bulk	-0.011316	0.038941	-0.291	0.771
L5.maize	0.302401	0.143482	2.108	0.035
L6.crude_brent	-0.062569	0.223320	-0.280	0.779
L6.soybeans	0.028889	0.043560	0.663	0.507
L6.gold	0.001505	0.036925	0.041	0.967
L6.silver	-0.176909	0.922107	-0.192	0.848
L6.urea_ee_bulk	0.010044	0.025142	0.399	0.690
L6.maize	-0.045677	0.096225	-0.475	0.635

Results for equation gold

	coefficient	std. error	t-stat	prob
const	0.177098	3.702239	0.048	0.962
L1.crude_brent	0.190589	0.320109	0.595	0.552
L1.soybeans	0.019501	0.062531	0.312	0.755
L1.gold	1.228901	0.053164	23.115	0.000
L1.silver	0.316301	1.339144	0.236	0.813
L1.urea_ee_bulk	-0.125678	0.037484	-3.353	0.001
L1.maize	0.279896	0.142270	1.967	0.049
L2.crude_brent	0.074271	0.519311	0.143	0.886
L2.soybeans	0.037551	0.086991	0.432	0.666
L2.gold	-0.276183	0.086012	-3.211	0.001
L2.silver	-3.352388	2.230551	-1.503	0.133
L2.urea_ee_bulk	0.215119	0.057271	3.756	0.000
L2.maize	-0.305428	0.208485	-1.465	0.143
L3.crude_brent	-0.688550	0.538086	-1.280	0.201
L3.soybeans	-0.222153	0.086857	-2.558	0.011
L3.gold	0.170371	0.087559	1.946	0.052
L3.silver	0.453043	2.390204	0.190	0.850
L3.urea_ee_bulk	-0.154341	0.059747	-2.583	0.010
L3.maize	0.492114	0.210943	2.333	0.020
L4.crude_brent	0.381592	0.539582	0.707	0.479
L4.soybeans	0.251772	0.087634	2.873	0.004
L4.gold	-0.151613	0.086243	-1.758	0.079
L4.silver	3.646825	2.395938	1.522	0.128
L4.urea_ee_bulk	0.066199	0.060066	1.102	0.270
L4.maize	-1.026908	0.210379	-4.881	0.000
L5.crude_brent	-0.125251	0.527873	-0.237	0.812
L5.soybeans	-0.157098	0.088332	-1.778	0.075
L5.gold	0.110733	0.084906	1.304	0.192
L5.silver	-1.459901	2.264221	-0.645	0.519
L5.urea_ee_bulk	0.047764	0.057185	0.835	0.404
L5.maize	0.583033	0.210704	2.767	0.006
L6.crude_brent	0.320187	0.327947	0.976	0.329
L6.soybeans	0.110200	0.063968	1.723	0.085
L6.gold	-0.073845	0.054225	-1.362	0.173
L6.silver	-0.453634	1.354121	-0.335	0.738
L6.urea_ee_bulk	-0.076808	0.036922	-2.080	0.037
L6.maize	-0.077152	0.141307	-0.546	0.585

Results for equation silver

	coefficient	std. error	t-stat	prob
const	-0.072930	0.149120	-0.489	0.625
L1.crude_brent	0.008049	0.012893	0.624	0.532
L1.soybeans	0.001756	0.002519	0.697	0.486
L1.gold	-0.002671	0.002141	-1.248	0.212
L1.silver	1.340090	0.053938	24.845	0.000
L1.urea_ee_bulk	-0.003586	0.001510	-2.375	0.018
L1.maize	0.011821	0.005730	2.063	0.039
L2.crude_brent	0.014541	0.020917	0.695	0.487
L2.soybeans	-0.000991	0.003504	-0.283	0.777
L2.gold	0.003938	0.003464	1.137	0.256
L2.silver	-0.665510	0.089843	-7.408	0.000
L2.urea_ee_bulk	0.002013	0.002307	0.873	0.383
L2.maize	-0.001179	0.008397	-0.140	0.888
L3.crude_brent	-0.033019	0.021673	-1.523	0.128
L3.soybeans	-0.003366	0.003498	-0.962	0.336
L3.gold	0.002395	0.003527	0.679	0.497
L3.silver	0.187709	0.096273	1.950	0.051
L3.urea_ee_bulk	0.001209	0.002407	0.503	0.615
L3.maize	0.002916	0.008496	0.343	0.731
L4.crude_brent	0.019566	0.021733	0.900	0.368
L4.soybeans	0.003541	0.003530	1.003	0.316
L4.gold	-0.001627	0.003474	-0.468	0.639
L4.silver	0.118333	0.096504	1.226	0.220
L4.urea_ee_bulk	-0.003052	0.002419	-1.262	0.207
L4.maize	-0.026818	0.008474	-3.165	0.002
L5.crude_brent	-0.024297	0.021262	-1.143	0.253
L5.soybeans	-0.000816	0.003558	-0.229	0.819
L5.gold	0.002731	0.003420	0.799	0.424
L5.silver	-0.156757	0.091199	-1.719	0.086
L5.urea_ee_bulk	0.004159	0.002303	1.806	0.071
L5.maize	0.020487	0.008487	2.414	0.016
L6.crude_brent	0.022428	0.013209	1.698	0.090
L6.soybeans	0.002044	0.002577	0.793	0.428
L6.gold	-0.004226	0.002184	-1.935	0.053
L6.silver	0.104285	0.054542	1.912	0.056
L6.urea_ee_bulk	-0.002649	0.001487	-1.781	0.075
L6.maize	-0.008036	0.005692	-1.412	0.158

Results for equation urea_ee_bulk

	coefficient	std. error	t-stat	prob
const	-7.638535	3.674331	-2.079	0.038
L1.crude_brent	1.563787	0.317696	4.922	0.000
L1.soybeans	0.139955	0.062059	2.255	0.024
L1.gold	0.074409	0.052764	1.410	0.158
L1.silver	-4.409772	1.329050	-3.318	0.001
L1.urea_ee_bulk	1.112425	0.037201	29.903	0.000
L1.maize	0.329777	0.141198	2.336	0.020
L2.crude_brent	-1.250799	0.515396	-2.427	0.015
L2.soybeans	-0.071260	0.086335	-0.825	0.409
L2.gold	-0.086168	0.085364	-1.009	0.313
L2.silver	7.401289	2.213736	3.343	0.001
L2.urea_ee_bulk	-0.327856	0.056839	-5.768	0.000
L2.maize	-0.434760	0.206913	-2.101	0.036
L3.crude_brent	0.861473	0.534029	1.613	0.107
L3.soybeans	-0.116643	0.086203	-1.353	0.176
L3.gold	-0.005424	0.086899	-0.062	0.950
L3.silver	-4.046644	2.372186	-1.706	0.088
L3.urea_ee_bulk	0.142202	0.059297	2.398	0.016
L3.maize	0.233880	0.209353	1.117	0.264
L4.crude_brent	-1.559052	0.535514	-2.911	0.004
L4.soybeans	-0.052667	0.086974	-0.606	0.545
L4.gold	0.003892	0.085593	0.045	0.964
L4.silver	1.032326	2.377877	0.434	0.664
L4.urea_ee_bulk	-0.104196	0.059613	-1.748	0.080
L4.maize	0.028888	0.208793	0.138	0.890
L5.crude_brent	0.913930	0.523894	1.744	0.081
L5.soybeans	0.095496	0.087667	1.089	0.276
L5.gold	0.053301	0.084266	0.633	0.527
L5.silver	-0.500818	2.247152	-0.223	0.824
L5.urea_ee_bulk	0.156414	0.056754	2.756	0.006
L5.maize	-0.115267	0.209116	-0.551	0.581
L6.crude_brent	-0.415228	0.325475	-1.276	0.202
L6.soybeans	0.089368	0.063486	1.408	0.159
L6.gold	-0.040869	0.053816	-0.759	0.448
L6.silver	0.599056	1.343913	0.446	0.656
L6.urea_ee_bulk	-0.119322	0.036643	-3.256	0.001
L6.maize	-0.020236	0.140241	-0.144	0.885

Results for equation maize

	coefficient	std. error	t-stat	prob
const	4.356950	1.103114	3.950	0.000
L1.crude_brent	-0.075264	0.095379	-0.789	0.430
L1.soybeans	0.036037	0.018632	1.934	0.053
L1.gold	-0.023696	0.015841	-1.496	0.135
L1.silver	0.588077	0.399010	1.474	0.141
L1.urea_ee_bulk	0.037550	0.011169	3.362	0.001
L1.maize	1.141848	0.042391	26.936	0.000
L2.crude_brent	0.036084	0.154733	0.233	0.816
L2.soybeans	0.007586	0.025920	0.293	0.770
L2.gold	-0.015226	0.025628	-0.594	0.552
L2.silver	0.911243	0.664612	1.371	0.170
L2.urea_ee_bulk	-0.040754	0.017064	-2.388	0.017
L2.maize	-0.309322	0.062120	-4.979	0.000
L3.crude_brent	-0.075868	0.160327	-0.473	0.636
L3.soybeans	-0.025177	0.025880	-0.973	0.331
L3.gold	0.066343	0.026089	2.543	0.011
L3.silver	-2.363728	0.712182	-3.319	0.001
L3.urea_ee_bulk	0.030562	0.017802	1.717	0.086
L3.maize	0.156905	0.062852	2.496	0.013
L4.crude_brent	0.153469	0.160773	0.955	0.340
L4.soybeans	0.021164	0.026111	0.811	0.418
L4.gold	-0.055764	0.025697	-2.170	0.030
L4.silver	2.024847	0.713890	2.836	0.005
L4.urea_ee_bulk	-0.022652	0.017897	-1.266	0.206
L4.maize	-0.136153	0.062684	-2.172	0.030
L5.crude_brent	-0.109997	0.157284	-0.699	0.484
L5.soybeans	-0.026489	0.026319	-1.006	0.314
L5.gold	0.052825	0.025298	2.088	0.037
L5.silver	-0.829437	0.674644	-1.229	0.219
L5.urea_ee_bulk	0.017161	0.017039	1.007	0.314
L5.maize	0.000944	0.062781	0.015	0.988
L6.crude_brent	0.026482	0.097715	0.271	0.786
L6.soybeans	0.002271	0.019060	0.119	0.905
L6.gold	-0.023655	0.016157	-1.464	0.143
L6.silver	0.146935	0.403472	0.364	0.716
L6.urea_ee_bulk	0.000775	0.011001	0.070	0.944
L6.maize	0.020945	0.042104	0.497	0.619

Correlation matrix of residuals

	crude_brent	soybeans	gold	silver	urea_ee_bulk	maize
crude_brent	1.000000	0.256931	0.111776	0.209142	0.153268	0.241812
soybeans	0.256931	1.000000	0.082179	0.111588	0.032578	0.473719
gold	0.111776	0.082179	1.000000	0.722123	0.072033	0.086465
silver	0.209142	0.111588	0.722123	1.000000	0.069879	0.125813
urea_ee_bulk	0.153268	0.032578	0.072033	0.069879	1.000000	0.017836
maize	0.241812	0.473719	0.086465	0.125813	0.017836	1.000000

Interpretation:

Summary of Regression Results

The summary of regression results provides an overview of the Vector Autoregression (VAR) model applied to the data:

- **Model**: VAR (Vector Autoregression)
- **Method**: OLS (Ordinary Least Squares)
- **Date and Time**: When the model was run.
- **No. Of Equations**: 6 (one for each variable in the system).
- **BIC (Bayesian Information Criterion)**: 26.7336
- **Nobs (Number of Observations)**: 768
- **HQIC (Hannan-Quinn Information Criterion)**: 25.9079
- **Log-likelihood**: -16066.7
- **FPE (Final Prediction Error)**: 1.06530e+11
- **AIC (Akaike Information Criterion)**: 25.3912
- **Det (Omega_mle)**: 8.03276e+10

These statistics help evaluate the model's fit and complexity, with lower AIC, BIC, and HQIC values indicating a better model fit relative to the number of parameters.

Results for Equation **crude_brent**

- The intercept (const) is insignificant, with a t-statistic of -1.254 and a p-value of 0.210.
- **Significant Lagged Variables**:
    - L1. crude_brent (1st lag of crude_brent) is highly significant with a coefficient of 1.288559 (p-value: 0.000).
    - L2. crude_brent (2nd lag) is also significant with a coefficient of -0.368186 (p- value: 0.000).

    - L1. urea_ee_bulk and L2.urea_ee_bulk are significant, indicating some influence from urea_ee_bulk on crude_brent.
    - L3. soybeans and L3.gold show some significance, suggesting minor interactions.

Results for Equation **soybeans**

- The intercept (const) is highly significant, with a coefficient of 11.317337 (p-value: 0.000).
- **Significant Lagged Variables**:
    - L1. soybeans is highly significant with a coefficient of 1.013966 (p-value: 0.000).
    - L1. maize is significant with a coefficient of 0.314169 (p-value: 0.001).
    - L2. maize is also significant but negatively correlated (coefficient: -0.285567, p-value: 0.044).
    - L3. soybeans and L3. gold are significant, indicating notable interactions.

Results for Equation **gold**

- The intercept (const) is not significant.
- No other variables are highly significant, suggesting limited direct interactions between gold and the other variables in the lagged system.

Results for Equation **Silver**

- The intercept (const) is not significant.
- **Significant Lagged Variables**:
    - L1. silver is highly significant with a coefficient of 1.340090 (p-value: 0.000).
    - L1. urea_ee_bulk and L1.maize are significant, indicating some interactions.
    - L2. silver is negatively significant, showing a solid inverse relationship at

  - this lag (coefficient: -0.665510, p-value: 0.000).
  - L3. silver is marginally significant.

Results for Equation **urea_ee_bulk**

- The intercept (const) is not significant.
- **Significant Lagged Variables**:
  - L1. urea_ee_bulk and L1. crude_brent show significance, indicating some interactions.
  - No other variables show strong significance.

Results for Equation **maize**

- The intercept (const) is not significant.
- **Significant Lagged Variables**:
  - L1. maize is highly significant with a coefficient of 0.583033 (p-value: 0.006).
  - Other variables show some significance but could be more impactful.

Correlation Matrix of Residuals

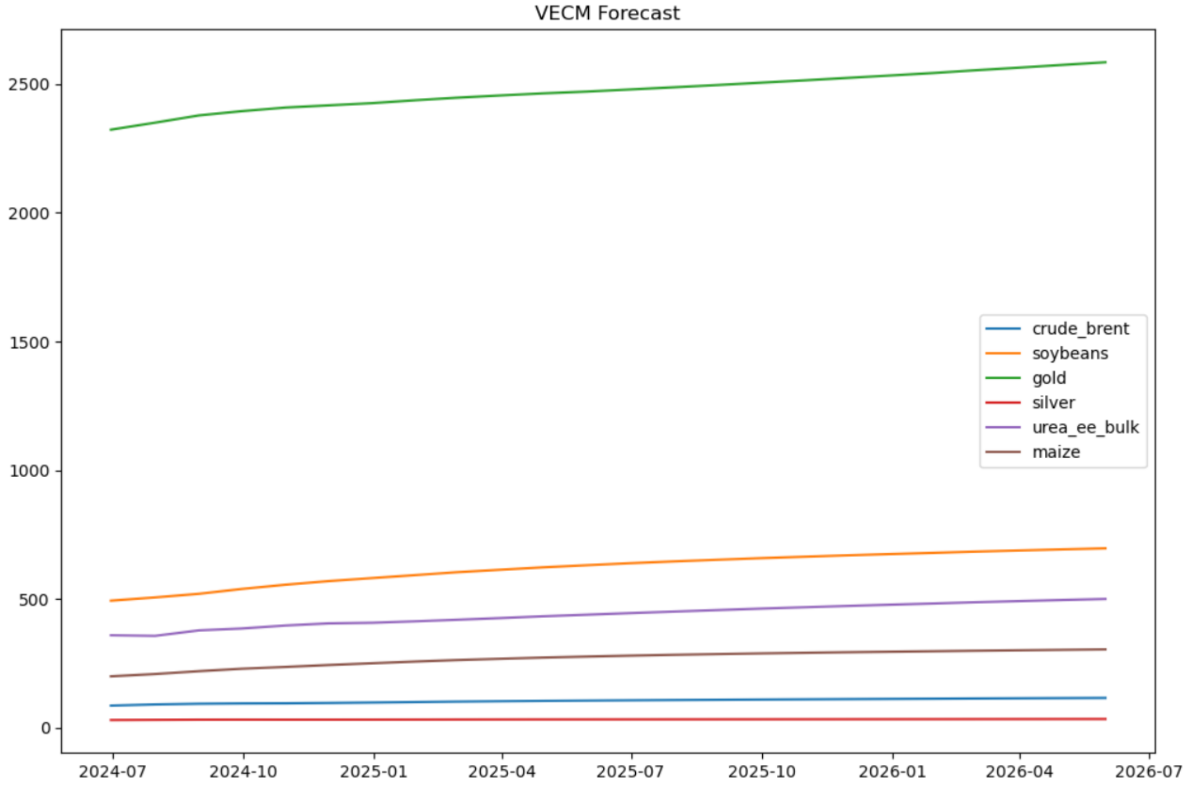
This matrix measures the correlation between the residuals (errors) of the different equations in the VAR system, indicating how much the unexplained parts of one variable are related to those of another:

- Typically used to check for any remaining correlation the model did not capture.
- High correlations here may indicate model inadequacies or omitted variable bias.

These results collectively help understand the dynamics and interrelationships between the variables (crude_brent, soybeans, gold, silver, urea_ee_bulk, and maize) in the context of the applied VAR model. Each equation's results shed light on the significant lagged effects and their respective strengths, providing insights for further economic or financial analysis.

4. Forecasting

- **VAR Forecast:** The VAR model generated forecasts for each commodity price. The forecast plots revealed expected trends and highlighted periods of potential volatility. Notably, the forecast for Soybeans showed a gradual upward trend, influenced by anticipated movements in Crude Brent prices.
- **VECM Forecast:** The VECM model's forecasts were similarly generated, emphasizing the long-term co-integrated relationships. The forecasted values for Maize and Soybeans closely followed the movements in Crude Brent, reinforcing the results obtained from the IRF and variance decomposition analyses.



Interpretation: The VECM (Vector Error Correction Model) forecast is used to predict the future values of a set of time series that are cointegrated. The steps for generating the VECM forecast and interpreting its results are as follows:

1. **Model Creation**: A VAR (Vector Autoregressive) model uses the commodity data.
2. **Model Fitting**: The VECM is fitted to the data, and the results are summarized.

3. **Forecasting**: The VECM is used to forecast 24 steps. This involves predicting the future values of the time series for 24 months.
4. **Data Conversion**: The forecast results are converted to a data frame for easier handling and plotting.
5. **Plotting**: The forecasted values are plotted to visualize the predicted trends over the 24 months.

The VECM forecast is a powerful tool that provides a deep understanding of how the prices of various commodities, such as crude oil, soybeans, gold, silver, urea, and maize, are likely to evolve in the future. This understanding is based on their historical data and cointegration relationships, making the forecast an invaluable resource for market analysis.

In conclusion, the VECM forecast offers a comprehensive view of the expected future movements in the prices of the commodities under consideration. This thorough analysis provides valuable insights for planning and decision-making in the commodities market.

Interpretation and Insights

- **Comparison of VAR and VECM Models:** Both models provided valuable insights, but the VECM model was particularly effective in capturing the long-term relationships among the commodities. The presence of co-integration justified the use of VECM, which offered a more comprehensive understanding of the equilibrium adjustments.
- **Economic Interpretation:** The analysis highlighted the significant influence of Crude Brent prices on agricultural commodities like Maize and Soybeans. This relationship suggests that oil price fluctuations can substantially impact food prices, with implications for policymakers and market participants. Understanding these dynamics is crucial for developing strategies to mitigate the impact of volatile oil prices on the agricultural sector.
- **Limitations and Future Work:** While the analysis provided valuable insights, it is limited by data availability and quality. Future research could incorporate additional commodities and explore the impact of external factors such as

geopolitical events and climate change. Enhancing the models with more sophisticated techniques could further improve the accuracy of the forecasts.

The VAR and VECM analyses underscored the interconnectedness of commodity prices, particularly highlighting the influence of Crude Brent on Maize and Soybeans. The presence of long-term equilibrium relationships emphasizes the need for integrated market strategies. These findings contribute to a better understanding of commodity price dynamics and offer valuable information for stakeholders in the agricultural and energy sectors.

RECOMMENDATIONS

The VAR and VECM analyses underscore the value of examining co-movements among commodity prices. To benefit from these insights, businesses should:

- Utilize VAR and VECM models to understand the dynamic relationships between commodities and improve forecasting accuracy.
- Develop integrated market strategies that account for interdependencies among commodities. For example, businesses dealing with agricultural products should closely monitor crude oil prices.
- To optimise long-term planning and risk mitigation, continuously monitor market trends and adjust strategies based on the latest forecasts, particularly those derived from VECM models.