QF621 Quantitative Trading Strategies

Pairs Trading: Application of Machine Learning and GARCH Model in the Foreign Exchange Market

Group Member:

CHEN Liangrui; JIANG Zihan; MAI Jialing; WONG Keat Ye; XUE Yuanhuang

CONTENT

INTRODUCTION

Pairs Trading in Forex

- Simultaneous long and short positions in two correlated currency pairs.
- Profits from the ***relative price movements*** rather than absolute price levels.
- ***Exploits*** mean reversion in correlated currency pairs.

Steps

1. Identify Pairs

2. Determine Spread

3. Trading Signal

4. Execute Trades

Project Overview

Data Source

- 35 historical Forex rates sourced from Yahoo Finance.
- Time frame: From 2014.01.01 to 2024.05.25

Pairs Selection

Method	*Introduction*
Pure PCA Analysis	• A ***dimensionality reduction technique*** that transforms a large set of variables into a smaller one that still contains ***most*** of the information. • It identifies the ***directions*** (principal components) in which the data varies the most.
PCA + OPTICS Clustering	• Combine PCA with OPTICS to ***leverage the strengths*** of both techniques for finding trading pairs. • OPTICS is a clustering algorithm that identifies clusters with ***varying densities***, which helps in finding groups of assets that ***behave similarly***.

Project Overview

Pairs Selection

Method	*Suitable for*	*Advantages*
Pure PCA Analysis	• Markets with ***high volatility*** • Markets with significant ***linear*** relationships • Markets with ***large data scales***	• Effectively captures the ***primary directions*** of variance
PCA + OPTICS Clustering	• ***Diversified*** markets • ***Complex*** markets with significant ***nonlinear*** relationships • Markets with ***strong local correlations***	• Identifies intricate market structures • Uncovers ***potential*** trading opportunities

Project Overview

Trading Signal

Method	*Introduction*
Traditional Methods	• Often involves ***simple statistical techniques*** to identify deviations from historical relationships between asset prices.
GARCH Model	• To model and ***forecast time-varying volatility*** in financial time series.

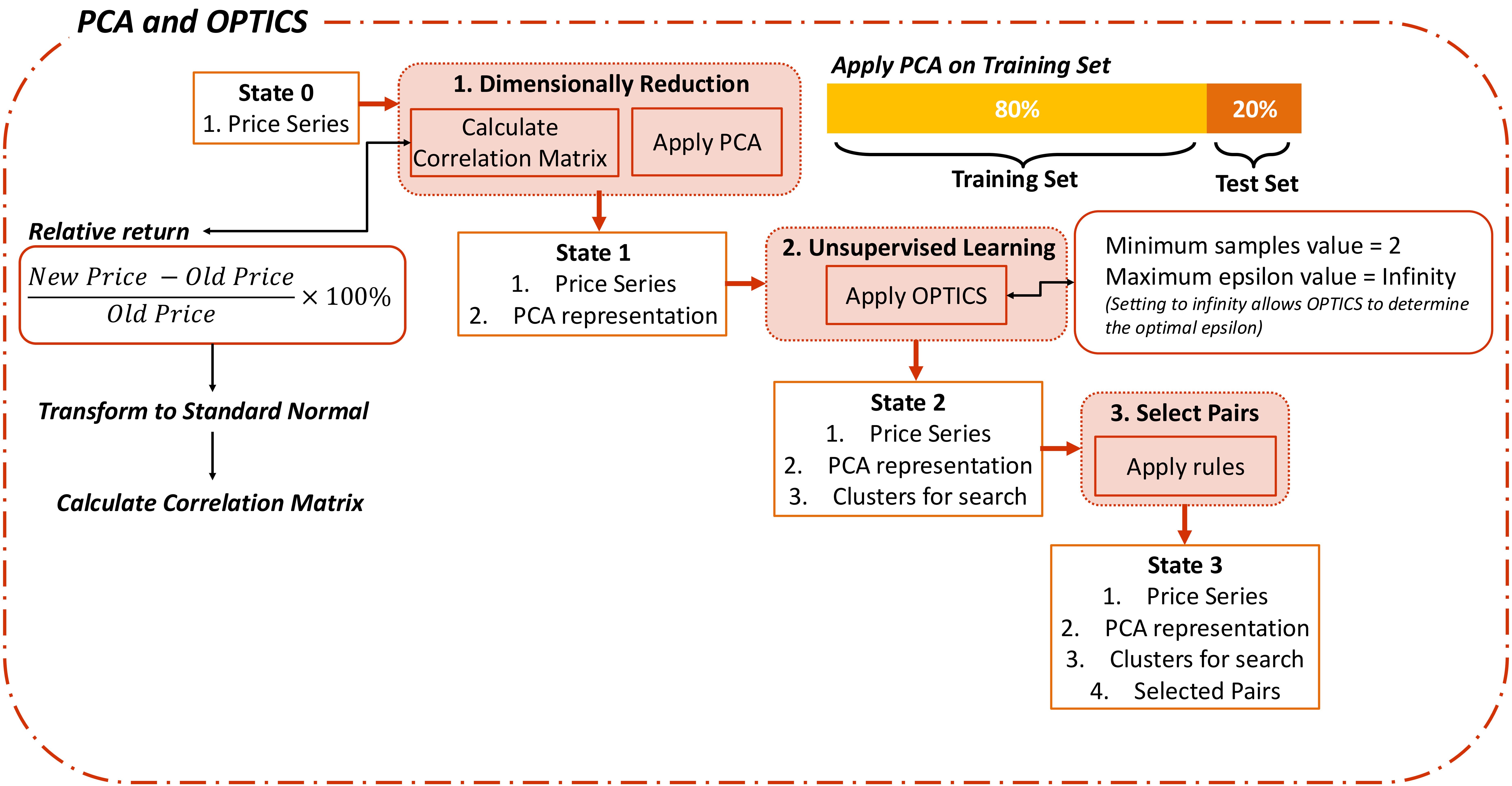
Project Overview

Trading Signal

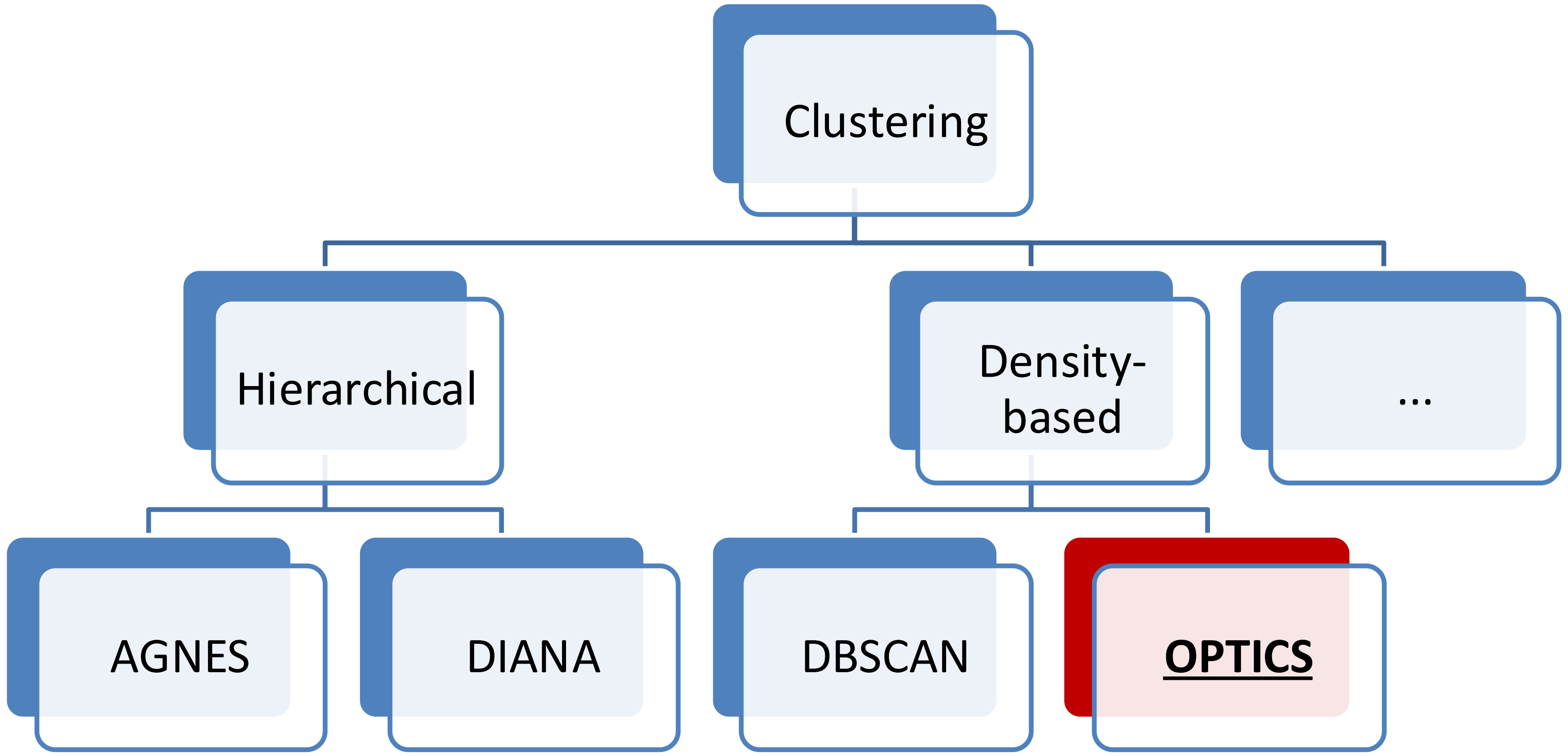
Method	*Suitable for*	*Advantages*
Traditional Methods	• Markets with ***stable*** volatility patterns • Markets with ***fewer data points***	• ***Simple*** and ***easy*** to implement • ***Quick*** computation • ***Less*** computational intensive
GARCH Model	• Markets with ***high and variable*** volatility • Markets with ***large*** datasets and **high-frequency** data • Markets where ***capturing volatility is crucial***	• Models time-varying volatility ***explicitly*** • ***More accurate*** spread estimation under conditions • Provides a ***more realistic*** measure of risk

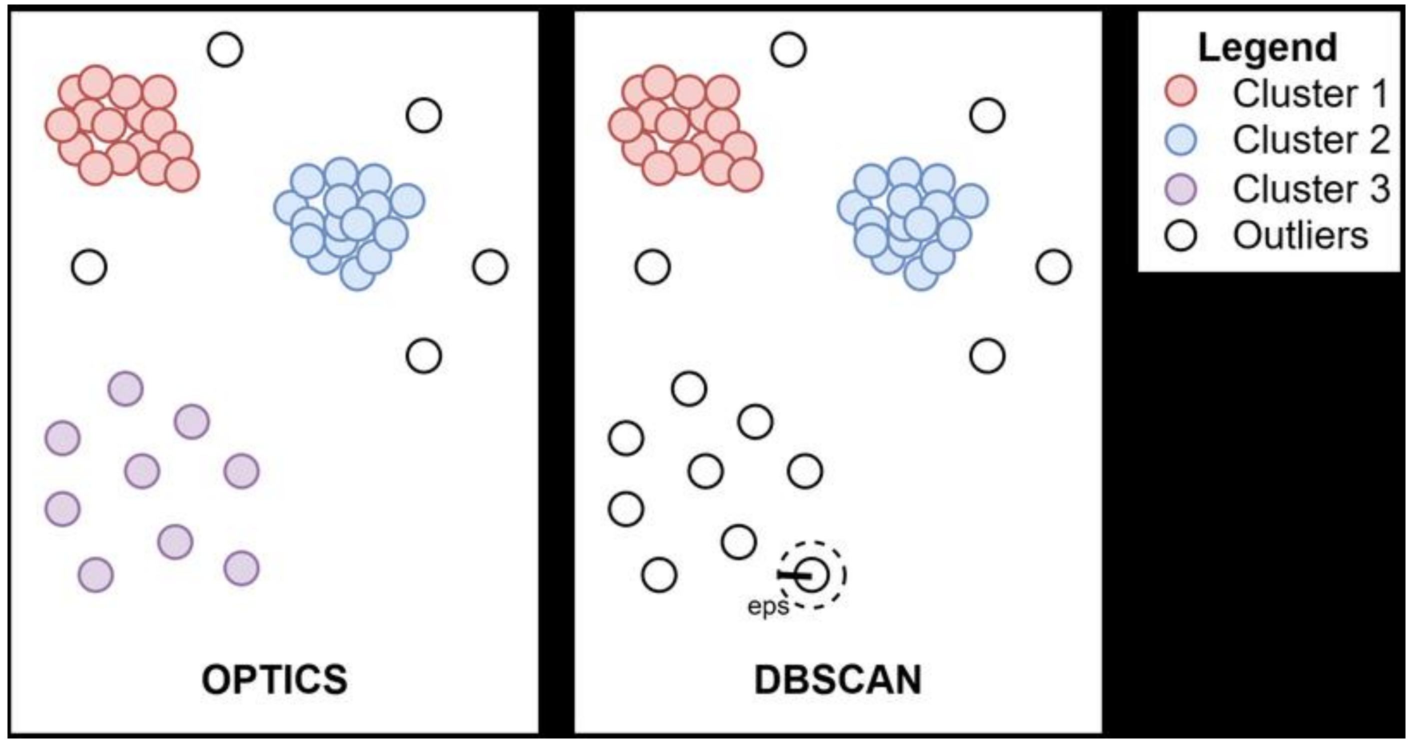
PAIRS SELECTION

Pairs Selection



What is OPTICS Clustering?





Python Package: https://scikit-learn.org/stable/modules/generated/sklearn.cluster.OPTICS.html

Pairs Selection

Check for Cointegration & Stationarity

Step 1 ***Identify the Most Influential Assets***

- Identify the assets with the highest and lowest absolute loading in the principal component.

Step 2 ***Log-Transform the Prices***

- To stabilize variance and convert multiplicative relationships into additive relationships.

Step 3 ***Perform OLS Regression***

- β: hedge ratio

Step 4 ***Perform Engle-Granger Cointegration Test***

- Reject if p-value > 0.5.

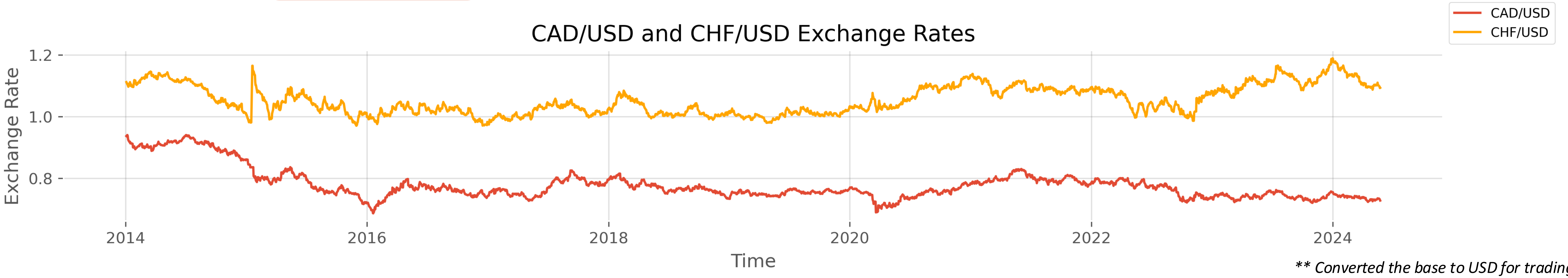
Step 5 ***Perform ADF Test for Stationarity***

- Performed on the residuals.
- Reject if p-value > 0.5.

Pairs Selection

PCA Results

CAD=X and CHF = X
(US Dollar / Canadian Dollar and US Dollar / Swiss Franc)



*** Converted the base to USD for trading*

Economic Intuition

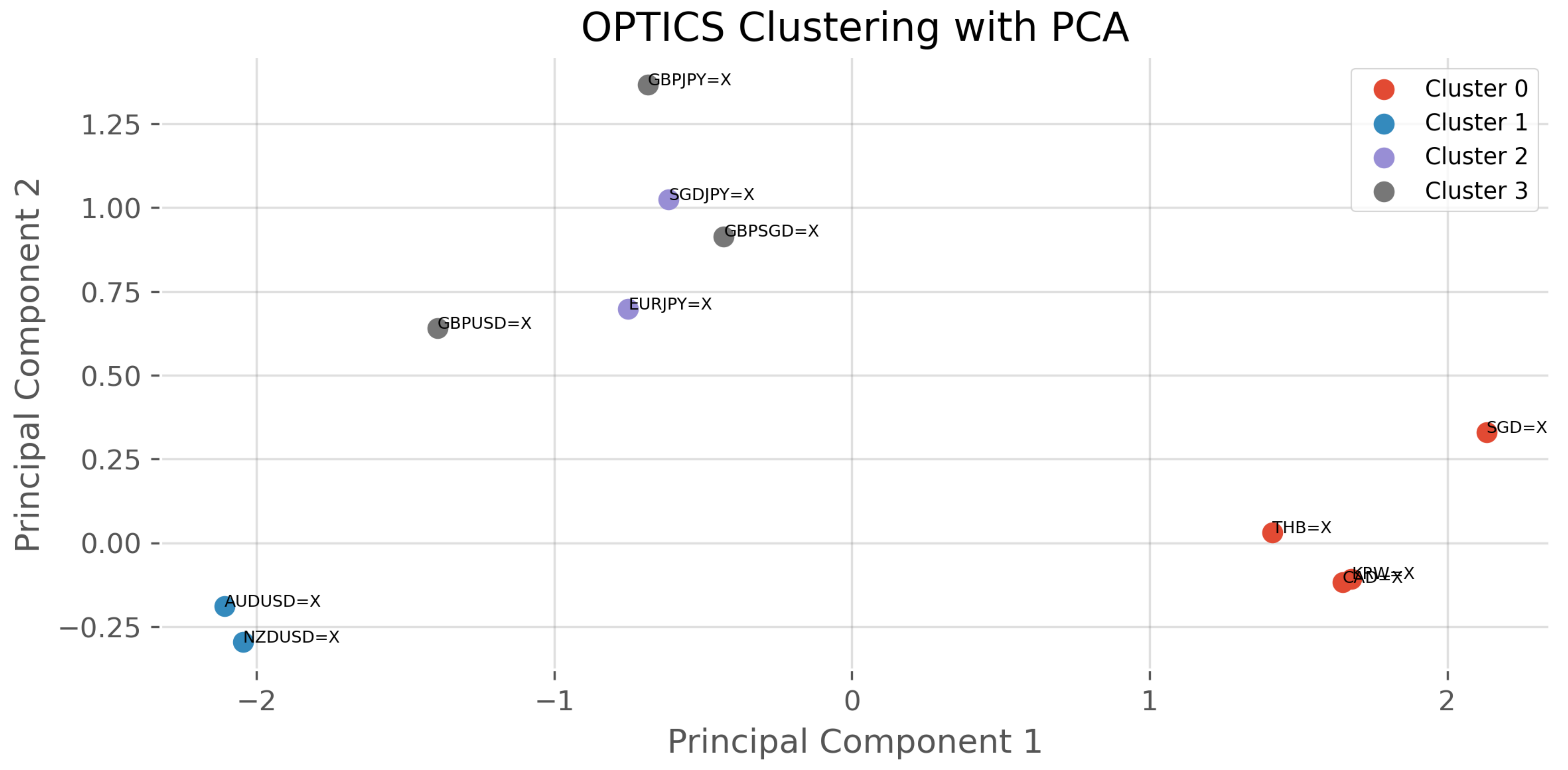
- **Different Economic Drivers:**
  1.CAD = X : Heavily influenced by the ***commodity market***, particularly oil.
  2.CHF = X: Often considered a ***safe-haven currency***.

- **Correlation and Diversification:**
  1.Correlation: Low correlation (***0.3160***) due to the different economic factors.
  - Opportunity for pairs trading, as the spread between them might offer mean-reverting characteristics.

  2.Diversification: The trader can potentially hedge against ***sector-specific risks***.

Pairs Selection

PCA + OPTICS Results



List of Noisy Points:
SGDMYR=X, EURSGD=X, SGDHKD=X, SGDIDR=X, SGDCNY=X, SGDTHB=X, SGDINR=X, SGDKRW=X, AUDSGD=X, NZDSGD=X, JPY=X, HKD=X, MYR=X, INR=X, CNY=X, PHP=X, IDR=X, CHF=X, MXN=X, VND=X, EURGBP=X, EURSEK=X, EURCHF=X, EURHUF=X

Cluster	Currency
0	SGD = X THB = X KRW = X CAD = X
1	AUDUSD = X NZDUSD = X
2	SGDJPY = X EURJPY = X
3	GBPSGD = X GBPJPY = X GBPUSD = X
Total	**11 pairs**

Pairs Selection

PCA + OPTICS Results

Cluster	Currency
0	SGD = X THB = X KRW = X CAD = X
1	AUDUSD = X NZDUSD = X
2	SGDJPY = X EURJPY = X
3	GBPSGD = X GBPJPY = X GBPUSD = X

Economic Explanation for the Clusters

Cluster 0: Based on ***USD pairings***, with significant trade relations with the US and influences from US economic data.

Cluster 1: Due to ***geographic proximity*** and ***similar economic structures***, heavily influenced by commodities and USD.

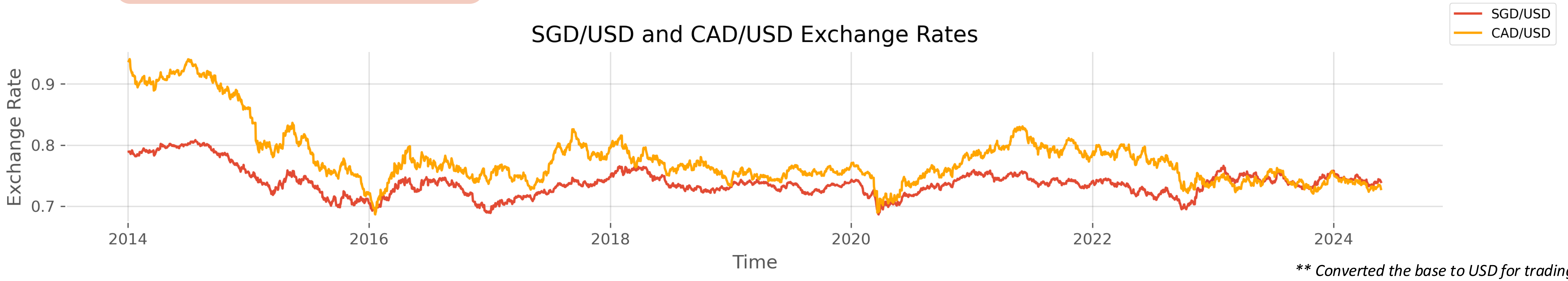
Cluster 2: Based on the common involvement of the ***Japanese Yen***, influenced by its safe-haven status and interest rate differentials.

Cluster 3: Due to the common involvement of the ***British Pound***, reflecting its global economic influence and diverse economic links.

PCA + OPTICS Results

SGD=X and CAD = X

(US Dollar / Singapore Dollar and US Dollar / Canadian Dollar)



*** Converted the base to USD for trading*

Economic Intuition

- **Economic Divergence:**
  1.SGD: A highly developed and open economy, heavily reliant on ***trade*** and ***services***.
  2.CAD: Heavily influenced by the ***commodity market***, particularly oil.

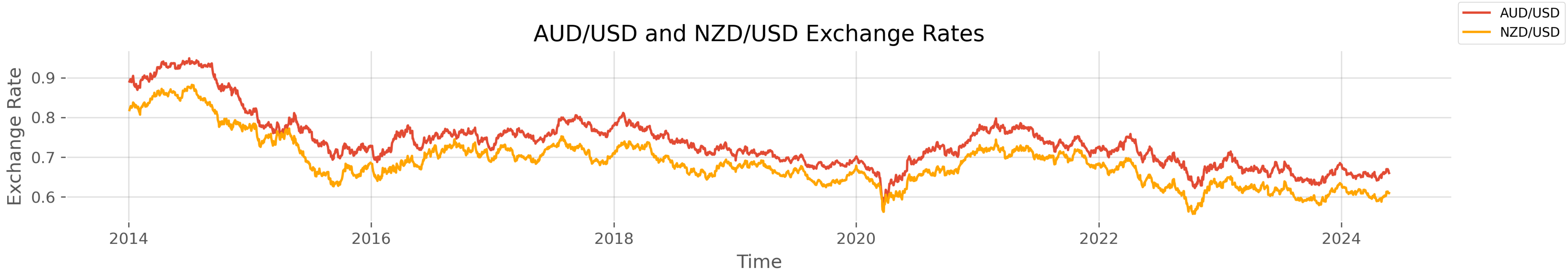
- **Interest Rate Differentials:**
  1.SGD: MAS uses the ***exchange rate*** as the main tool of monetary policy rather than interest rates.
  - Interest rates in Singapore are influenced by ***global interest rates*** and the ***domestic economy's performance***.

  2.CAD: Bank of Canada's ***interest rate decisions*** directly impact the CAD.
  - Higher interest rates in Canada attract foreign investment, strengthening the CAD, while lower rates can weaken it.

PCA + OPTICS Results

AUDUSD=X and NZDUSD = X

(Australian Dollar / US Dollar and New Zealand Dollar / US Dollar)



Economic Intuition

- **Economic Structures:**
  1.Australia: A ***diverse economy*** with key sectors including mining, agriculture, services, and manufacturing.
  2.New Zealand: A ***smaller, open economy*** heavily reliant on agriculture and dairy products.

- **Geopolitical Factors:**
  - Both countries have relatively ***stable*** political environments, but they can be affected by global geopolitical events that impact trade, commodity prices, and investor sentiment.

TRADING SIGNAL

Trading Signal

Traditional Method

1 Generate Spread Series

From OLS Regression

Intercept Coefficient

$$MSPREAD = Y_t - (a + bX_t)$$

Log price series of 2 assets

Decentralization

2 Calculation of Mean and Standard Deviation

$$\mu = \frac{1}{N}\sum_{t=1}^{N} MSPREAD_t$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(MSPREAD_t - \mu)^2}$$

3 Construction of Trading Signals

$$Buy\ Signal = MSPREAD_t < \mu - k\sigma$$

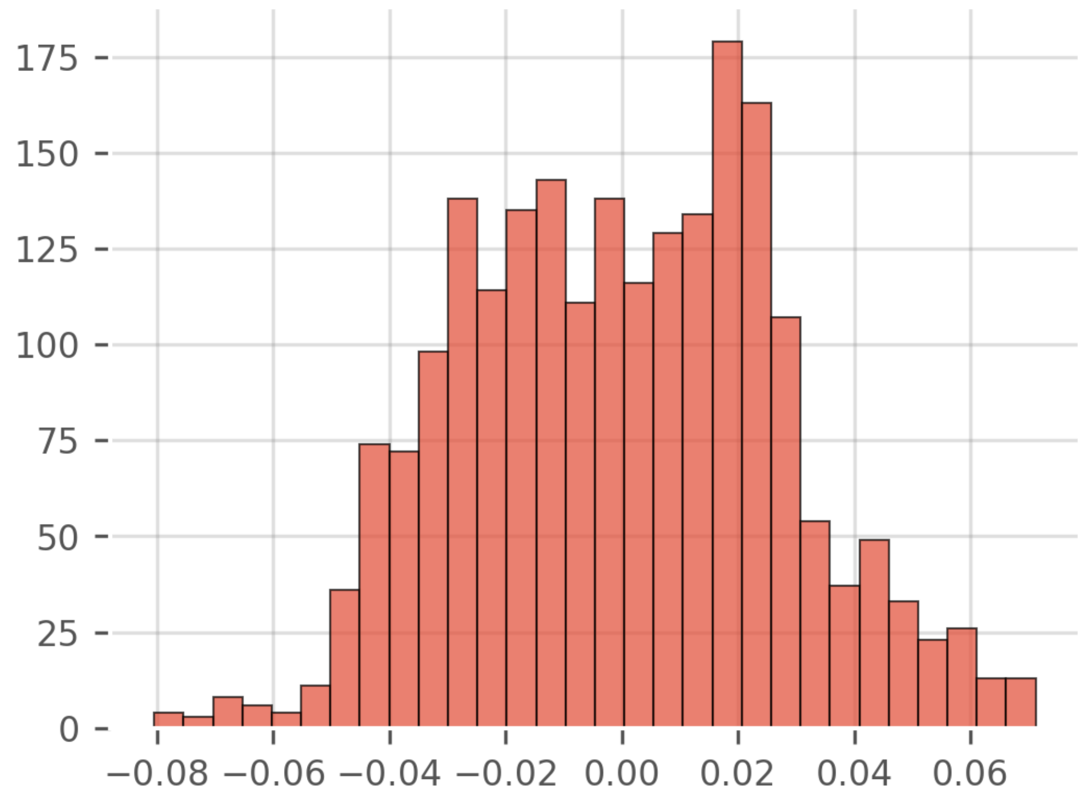
$$Sell\ Signal = MSPREAD_t > \mu + k\sigma$$

In our project, we set $\boldsymbol{k = 1}$.

Traditional Method

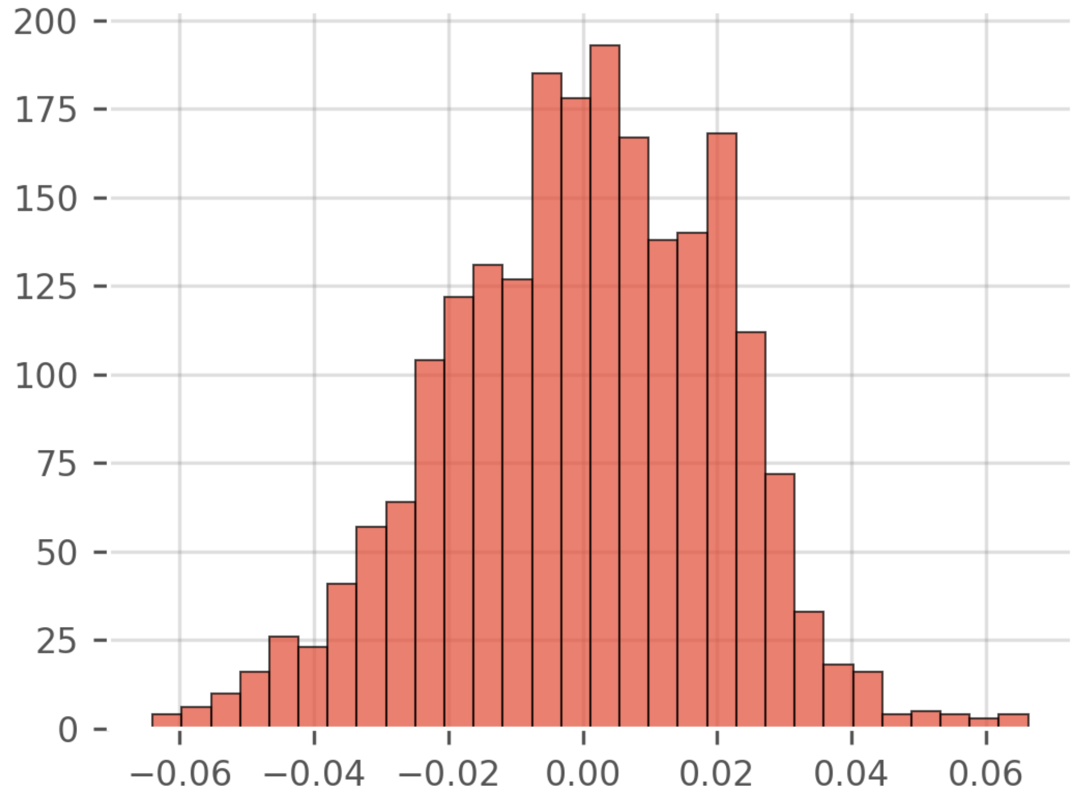
Descriptive Statistics of Decentralized Spread Series

SGD = X and CAD = X



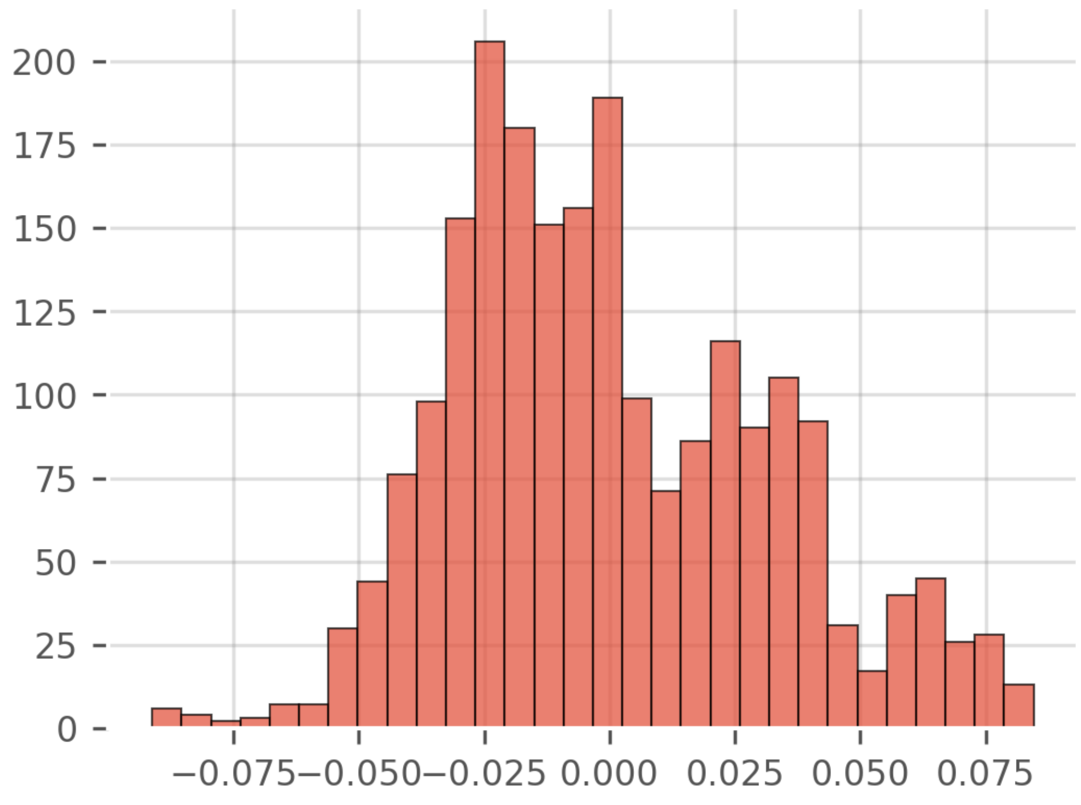
Mean = 0
Std. Dev. = 0.0272
Skewness = 0.0584
Kurtosis = 2.5515
Jarque-Bera Stats. : 19.4320
P-value (JB Test): 0

AUDUSD = X and NZDUSD = X



Mean = 0
Std. Dev. = 0.0205
Skewness = -0.1962
Kurtosis = 2.9459
Jarque-Bera Stats. : 14.1902
P-value (JB Test): 0.0008

CAD = X and CHF = X



Mean = 0
Std. Dev. = 0.0318
Skewness = 0.4245
Kurtosis = 2.7480
Jarque-Bera Stats. : 70.9398
P-value (JB Test): 0.0000

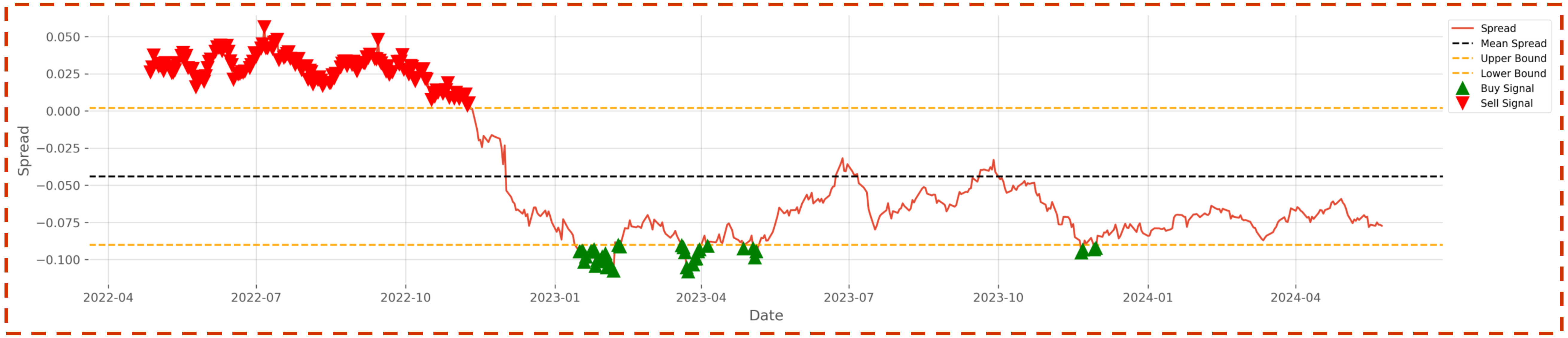
Trading Signal

Traditional Method

Construction of Trading Signals

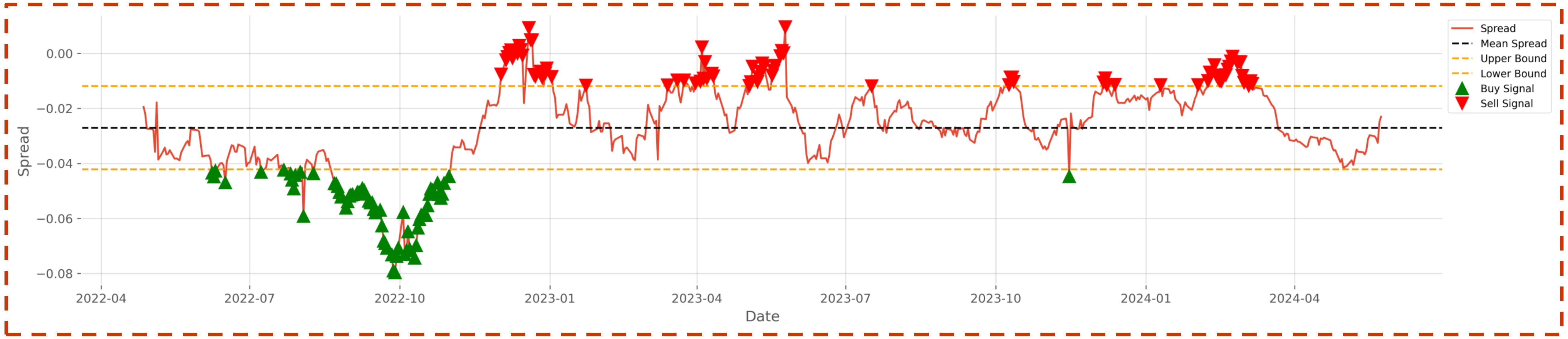
SGD = X and CAD = X

Holding: 365
Sell: 141
Buy: 37



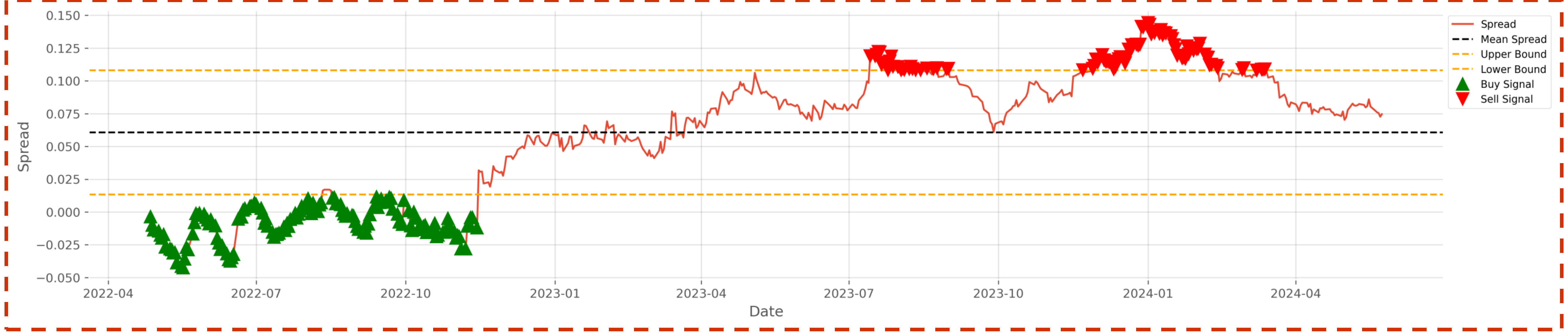
AUDUSD = X and NZDUSD = X

Holding: 400
Sell: 78
Buy: 65



CAD = X and CHF = X

Holding: 316
Sell: 140
Buy: 87



Trading Signal

GARCH Model Method

1 Fitting an AR(2) Model

$$MSPREAD_t = \underbrace{c}_{\text{Constant}} + \sum_{i=1}^{p} \underbrace{\phi_i}_{\text{Lag coefficients}} MSPREAD_{t-1} + \epsilon_t$$

2 Fitting a GARCH Model to AR(2) residuals

$$\epsilon_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2$$

Forecast σ^2

3 Adjust the spread

$$Adj.SPREAD_t = MSPREAD_t - c - \sum_{i=1}^{p} \phi_i MSPREAD_{t-1}$$

4 Construction of Trading Signals

$$Buy\ Signal = Adj.SPREAD_t < -k\sigma$$

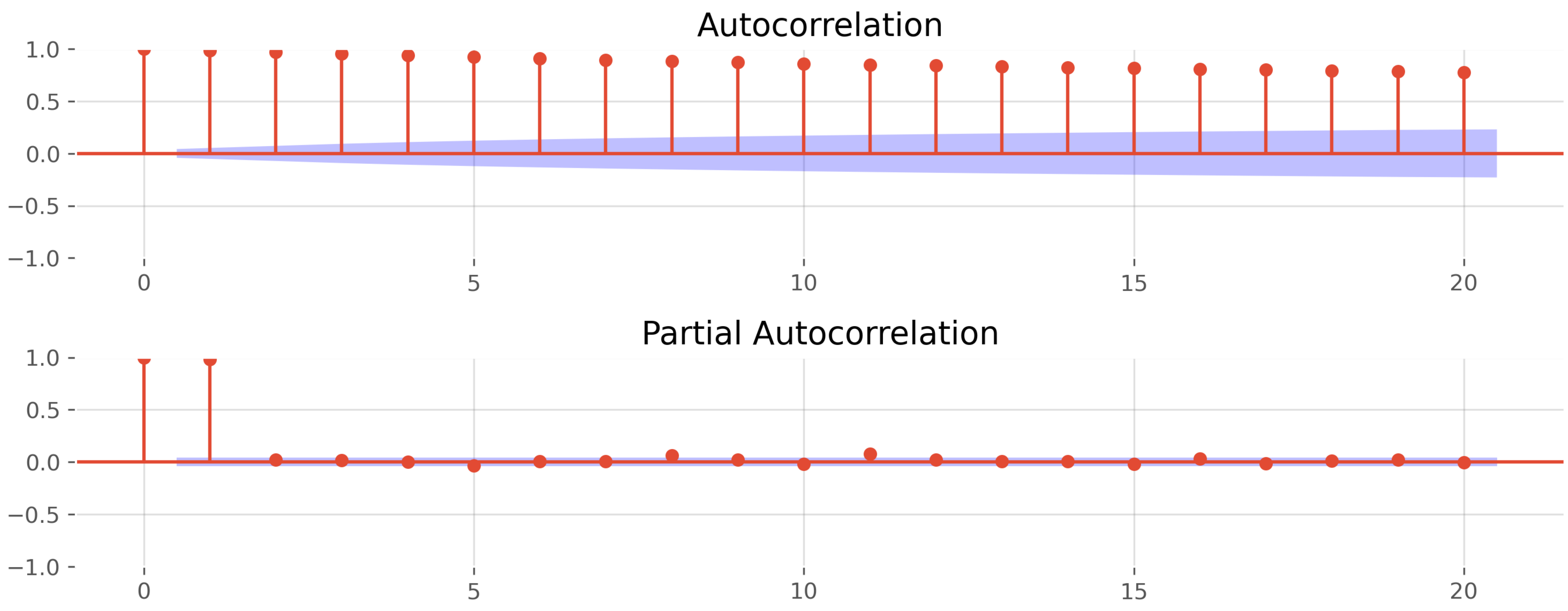
$$Sell\ Signal = Adj.SPREAD_t > k\sigma$$

In our project, we set $k = 1$.

Trading Signal

GARCH Model Method

Autocorrelation Test (All the pairs have the same results.)



PACF Significant spikes at lag 1 and 2 and then fluctuates around zero.

ACF A slow decay.

$\Bigg\}$ ***AR(2)***

Trading Signal

GARCH Model Method

ARCH-Lagrange Multiplier Test

- To indicate evidence of heteroskedasticity in the residuals of the GARCH(1,1) model.
  - P-value > 0.05 : Homoskedasticity
  - P-value < 0.05 : Heteroskedasticity

SGD=X and CAD=X

Lagrange Multiplier Test Statistic: 24.1818
p-value: 0.0071

AUDUSD = X and NZDUSD = X

Lagrange Multiplier Test Statistic: 30.9759
p-value: 0.0006

CAD = X and CHF = X

Lagrange Multiplier Test Statistic: 42.6111
p-value: 0.0

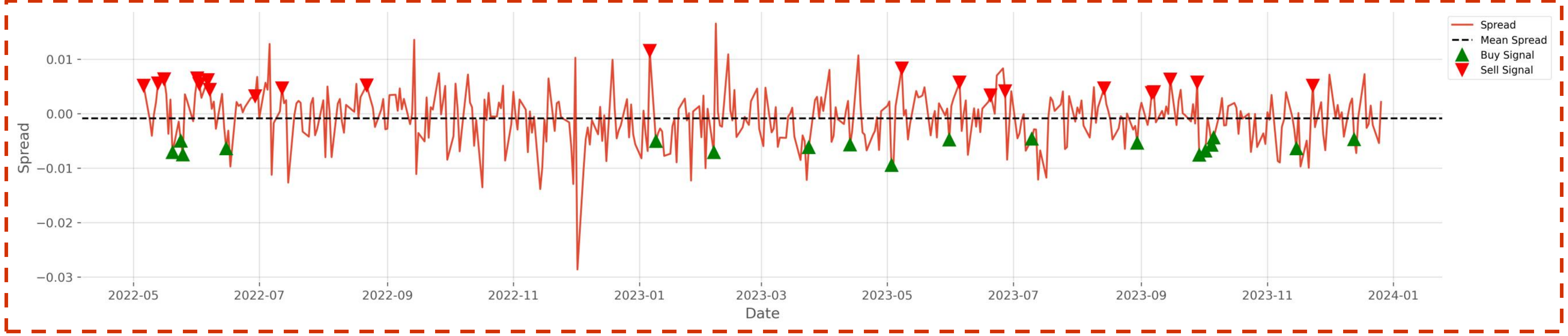
Heteroskedasticity

Trading Signal

GARCH Model Method Construction of Trading Signals

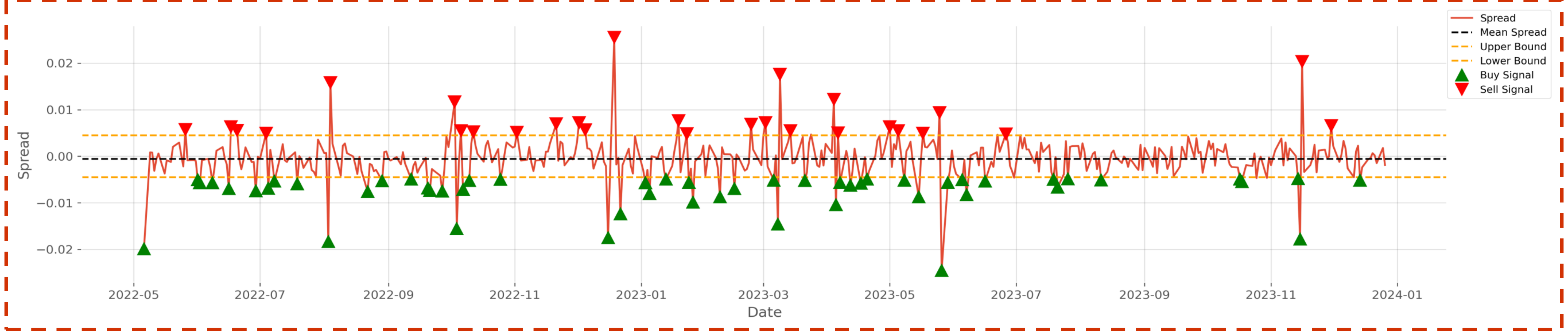
SGD = X and CAD = X

Holding: 315
Sell: 38
Buy: 75



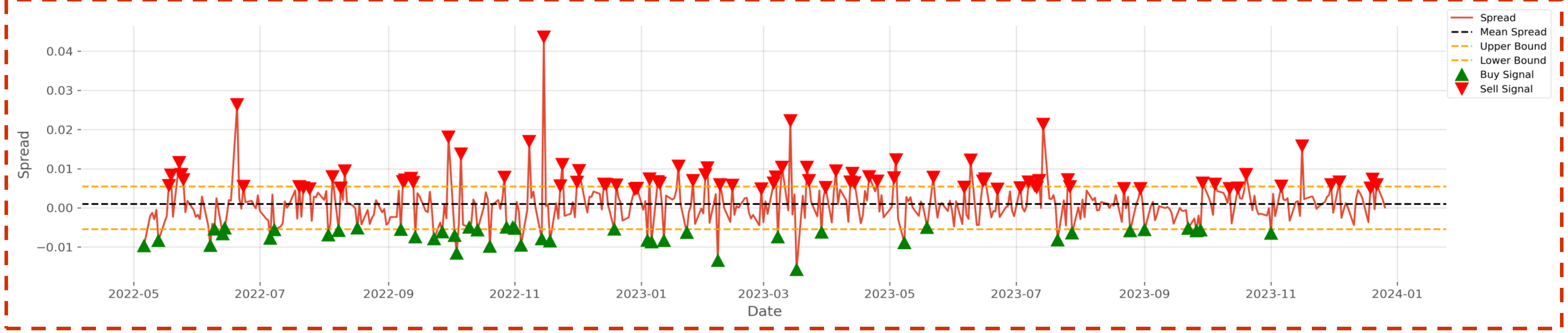
AUDUSD = X and NZDUSD = X

Holding: 347
Sell: 28
Buy: 53



CAD = X and CHF = X

Holding: 299
Sell: 84
Buy: 45



BACK-TESTING RESULT

Leverage

When it is ***Buy Signal*** -> Buy β spread and vice versa.

Where β is the OLS estimation coefficient

Back Testing Metrics

Cumulative Return	Cumulative of daily return
Sharpe Ratio	$\sqrt{252} \times \frac{\text{Average } Value\ of\ Daily\ Return}{Standard\ Deviation\ of\ Daily\ Return}$
Sortino Ratio	$\sqrt{252} \times \frac{\text{Average value } of\ daily\ return}{\sqrt{\frac{\sum Daily\ Return^2}{Number\ of\ Days}}}$
Max Drawdown (%)	$max\left(\frac{Maximum\ Drawdown\ (\$)}{Cumulative\ Maximum\ portfolio\ value} \times 100\%\right)$

Back-testing Result

Methods	Pairs	Cumulative Return (%)	Sharpe Ratio	Sortino Ratio
Traditional	SGD = X and CAD = X	7.55	0.54	0.87
	AUDUSD = X and NZDUSD = X	6.87	0.5	0.71
	CAD = X and CHF = X	7.64	0.43	0.78
GARCH Model	SGD = X and CAD = X	3.56	0.15	0.25
	AUDUSD = X and NZDUSD = X	37.97	0.61	0.96
	CAD = X and CHF = X	13.29	0.57	0.97

Back-testing Result

Key Observations

1. **Superior Performance of GARCH Model**
   - This suggests that the GARCH model is more effective in capturing and utilizing volatility information for generating trading signals.

2. **Enhanced Risk Management**
   - Lower drawdown reflects the GARCH model's ability to adapt to changing market conditions and mitigate significant losses during adverse periods.

3. **Consistency Across Pairs**
   - The performance improvement with the GARCH model was consistent across all tested currency pairs.
   - This consistency enhances the credibility of the GARCH model as a robust tool for pairs trading strategies across different forex markets.

4. **Weak Performance of PCA Pair**
   - Weak performance for the CAD=X and CHF=X pair in both methods, with smallest Sharpe and Sortino ratios.
   - The poor performance of the PCA pair can be attributed to its inability to effectively capture complex, non-linear relationships and varying densities within the data.

DISCUSSION

Discussion

Weakness of the Technique

1 High Computational Requirements

- The GARCH model is computationally intensive, requiring significant processing power and advanced technical expertise.

2 Data Dependency

- The strategy relies heavily on historical data to model volatility and generate trading signals.
- This dependency can be a limitation as past performance does not always predict future results, especially in rapidly changing markets.

3 Market Assumptions

- Pairs trading strategies, including those using the GARCH model, often assume that the price relationship between pairs will revert to the mean.
- If this assumption does not hold true, especially during prolonged trends or structural market changes, the strategy may incur losses.

Areas for Further Research

1 Incorporation of Macro-Economic Indicators

- Understanding the broader economic context may help in anticipating market movements more accurately.

2 Sentiment Analysis

- Incorporating sentiment analysis from news articles, social media platforms, and financial forums could provide additional insights into market sentiment.
- This can help in capturing market moods and trends that are not reflected in historical price data alone.

3 Regime Switching Models

- Combining the GARCH model with regime-switching models could help in identifying different market regimes (e.g., bull, bear, high volatility, low volatility).
- Adapting trading strategies based on the identified regime can improve performance and risk management.

4 High-Frequency Data

- Exploring the application of the GARCH model to high-frequency data can provide opportunities for intraday trading.

THANK

YOU