QF621 Quantitative Trading Strategies (AY2021/2022)

Research Project: Pairs Trading Strategy Using Machine Learning

Prepared By:
Brandon Lee
Dani Surya Pangestu
Gabriel Woon
Chan Pak Lam
Peter Chettiar
Wong Yong Wen

Table of Contents

1. Introduction

Commodity futures tend to have mean-reverting properties due to the push-pull effects of supply and demand. Traders have sought to profit from this relationship, especially with futures pair. The spot prices of these futures have a high degree of volatility hence traders take positions in the spreads between products instead which results in ever rebalancing market conditions which are perfect for pairs trading.

Hence, prices tend to move in relative lockstep with each other due to their substitutive and competing characteristics. Short-term deviations always mean revert due to fundamental drivers unless markets are completely disrupted. By identifying strongly cointegrated product pairs, it is possible to profit from a pair trading strategy.

2. Methodology

This research report seeks to implement machine learning in the entire process of the pairs trading framework and answer the following questions:

- Does it outperform the rational/fundamental selection process?
- Do the more complex algorithms produce better results?

The universe of 66 most liquid commodities is extracted from Reuters (Refinitiv). The full list of future products used can be found in the Appendix. The number of pairs in this universe will be 2,145.

3. Pairs Selection

As the popularity of pairs trading grows, it is increasingly harder to find rewarding pairs. The simplest procedure commonly applied is to generate all possible candidate pairs by considering the combination of every security with every other security in the dataset.

Two different problems arise immediately:

1. First, the computational cost of testing mean-reversion for all the possible combinations increases drastically as more securities are considered.
2. The second emerging problem is frequent when performing multiple hypothesis tests at once and is referred to as the multiple comparisons problem. If 100 hypothesis tests are performed (with a confidence level of 5%) the results should contain a false positive rate of 5%.

This problem was tackled by (Harlacher 2016), who found that Bonferroni correction leads to a very conservative selection of pairs and impedes the discovery of even truly cointegrated combinations. The paper recommends the effective pre-partitioning of the considered asset universe to reduce the number of feasible combinations and, therefore, the number of statistical tests. This aspect might lead the investor to pursue the usual, more restrictive approach of comparing securities only within the same sector.

This dramatically reduces the number of necessary statistical tests, consequently decreasing the likelihood of finding spurious relations. However, the simplicity of this process might also turn out to be a disadvantage. The more traders are aware of the pairs, the harder it is to find pairs not yet being traded in large volumes, leaving a smaller margin for profit.

This dilemma motivated the work of (Sarmento and Horta 2020) in the search for a methodology that lies in between these two scenes: an effective pre-partitioning of the universe of assets that does not limit the combination of pairs to relatively obvious solutions, while not enforcing excessive search combinations.

3.1. Unsupervised Machine Learning – Clustering

Unsupervised learning clustering techniques can solve the task at hand. The 3 clustering models selected are K-Means, Hierarchal and Affinity Propagation. The best-performing model is then selected based on its silhouette score. The timeframe for training the clustering models will be from 1/1/2009 to 31/12/2017.

K-Means Clustering

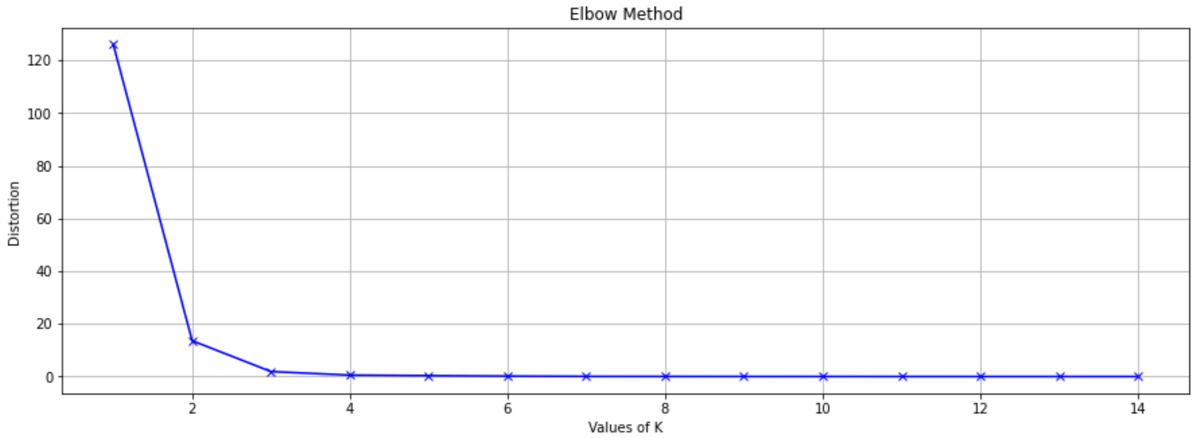


Fig 3.1.1 K-Means Clustering Elbow Method

Optimal number of clusters: 3

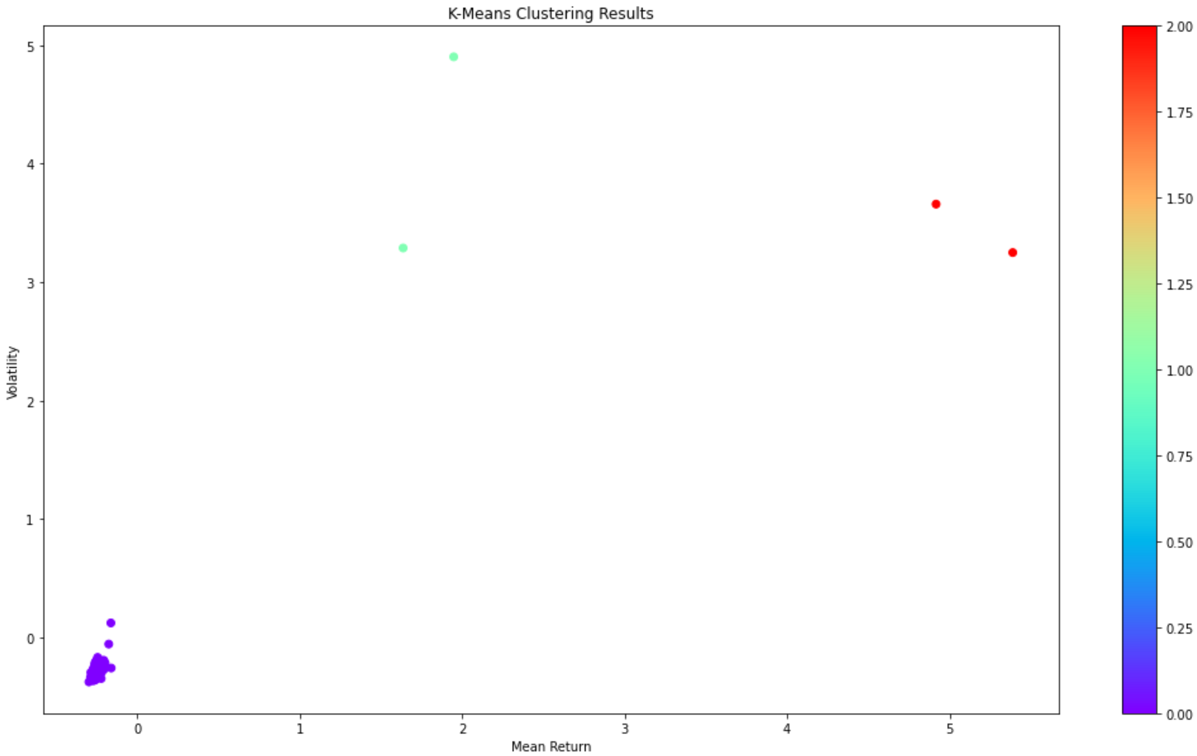


Fig 3.1.2 K-Means Clustering Plot

Hierarchal Clustering

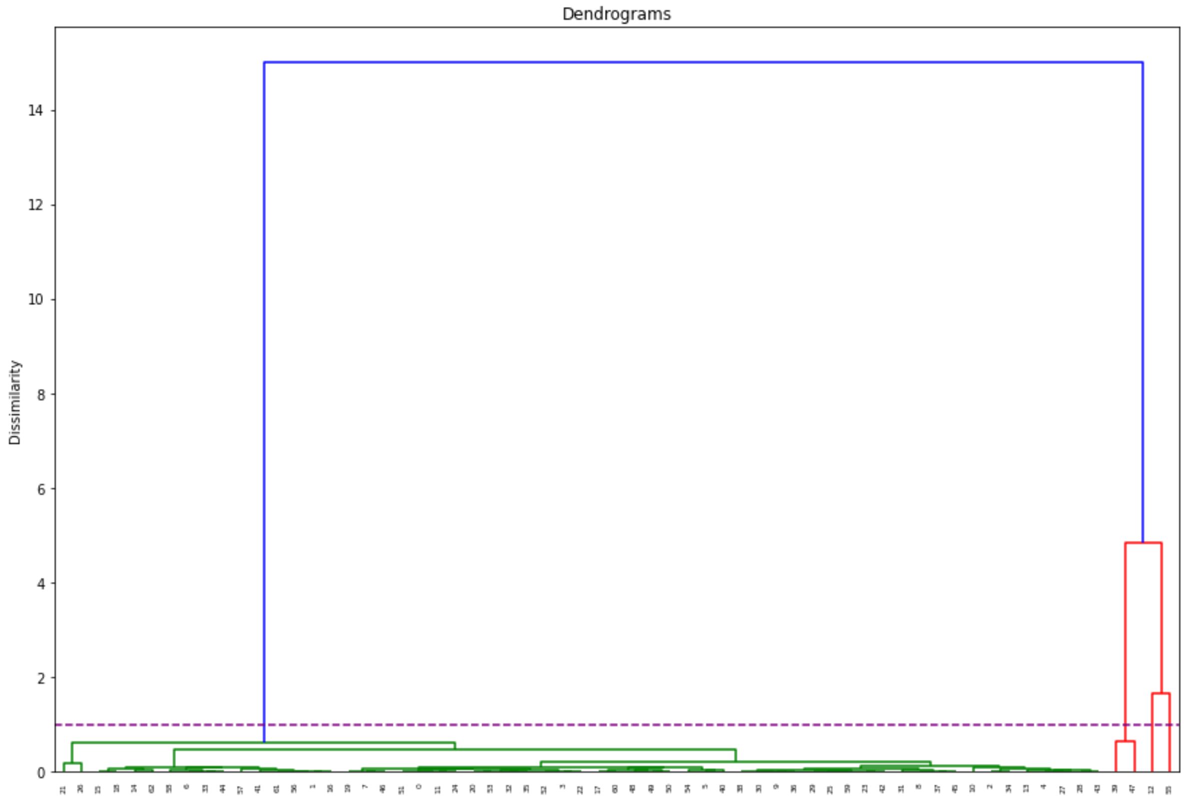


Fig 3.1.3 Hierarchal Clustering Dendrogram

Optimal number of clusters: 4

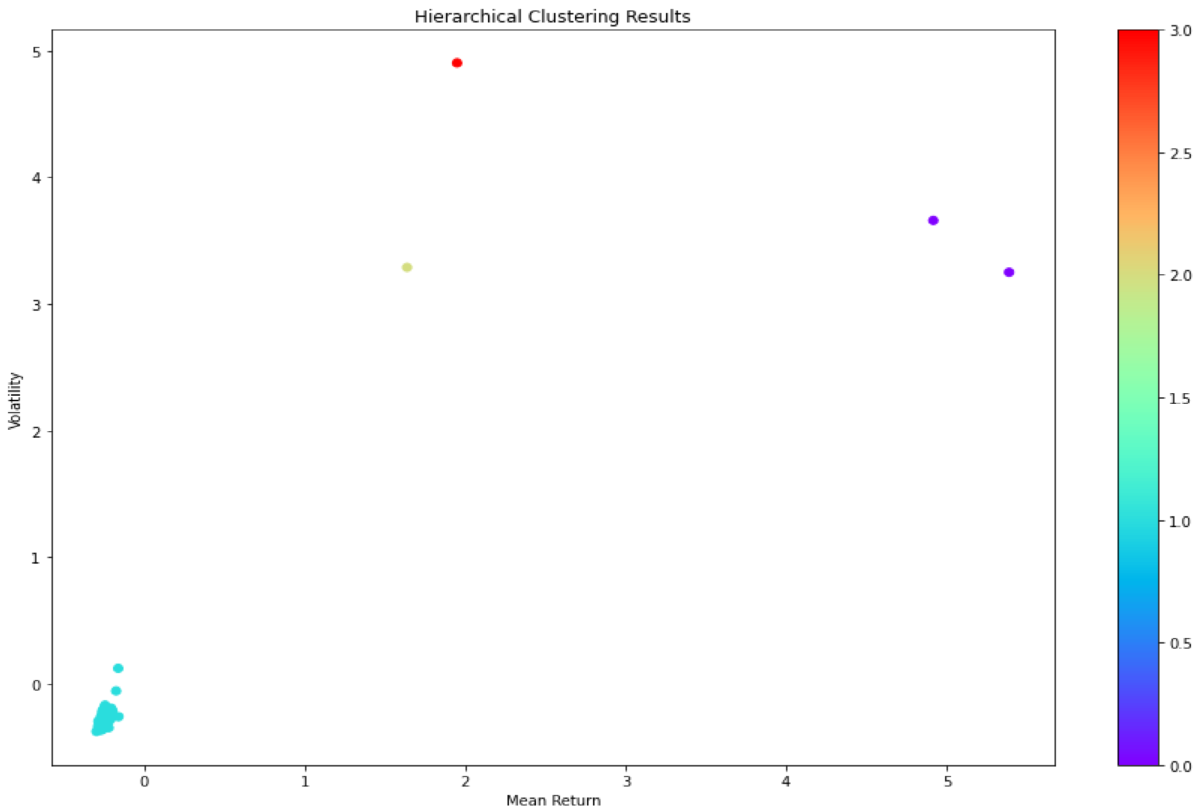


Fig 3.1.3 Hierarchal Clustering Plot

Affinity Propagation Clustering

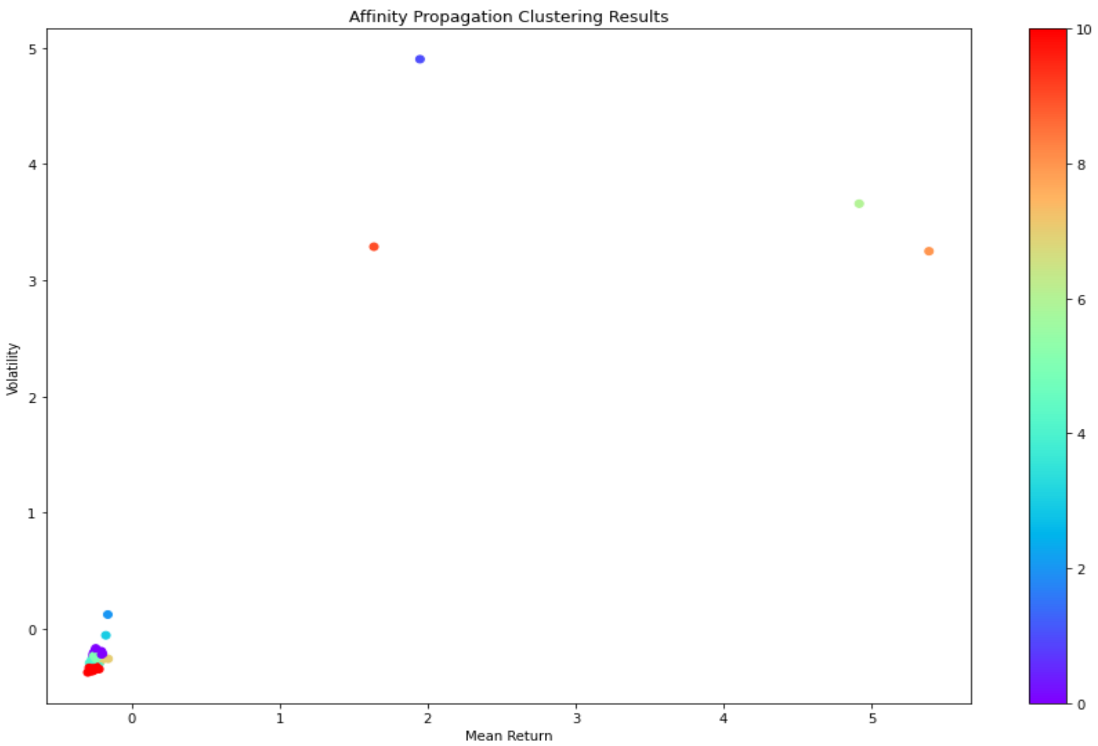


Fig 3.1.5 Affinity Propagation Clustering Plot

Cluster Evaluation

Silhouette Score:

- K-Means Clustering 0.29743
- Hierarchical Clustering 0.30728
- Affinity Propagation Clustering 0.94311

The clustering process resulted in 5 clusters with 662 unique pairs.

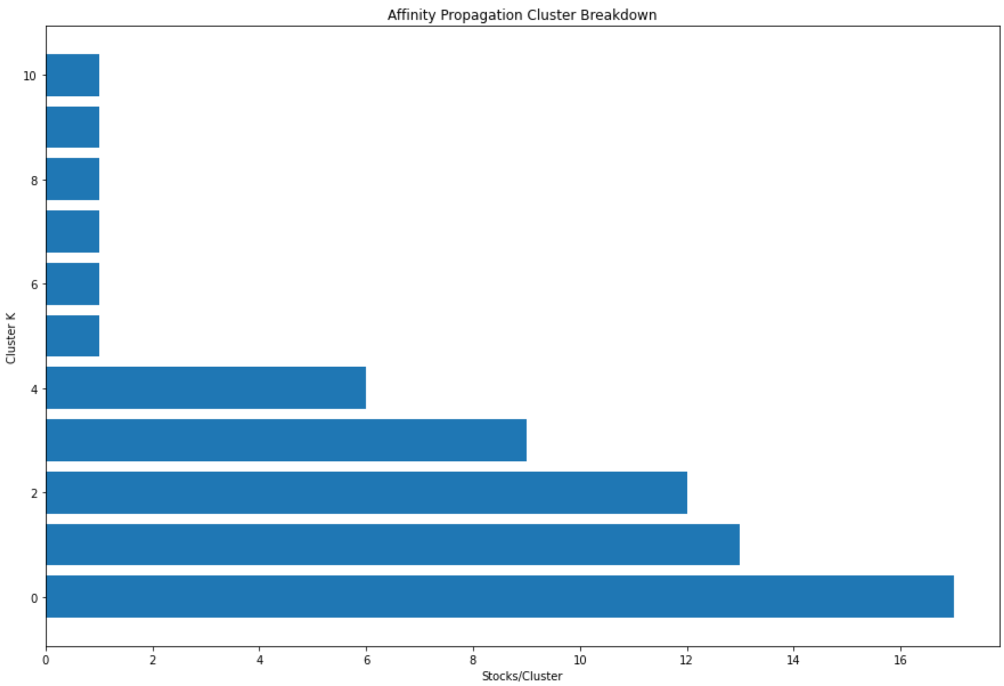


Fig 3.1.6 Affinity Propagation Cluster Breakdown

3.2. Absolute Rules of Disqualification (ARODs)

The classical pairs selection approach encompasses two steps:

1. Finding the appropriate candidate pairs
2. Selecting the most promising ones

Having generated the clusters of assets in the previous steps, it is still necessary to define a set of conditions for selecting the pairs to trade. The most common approaches to select pairs are the distance, cointegration, and correlation approaches. The cointegration approach was selected because of its simplicity and it performs better than the minimum distance and correlation approaches.

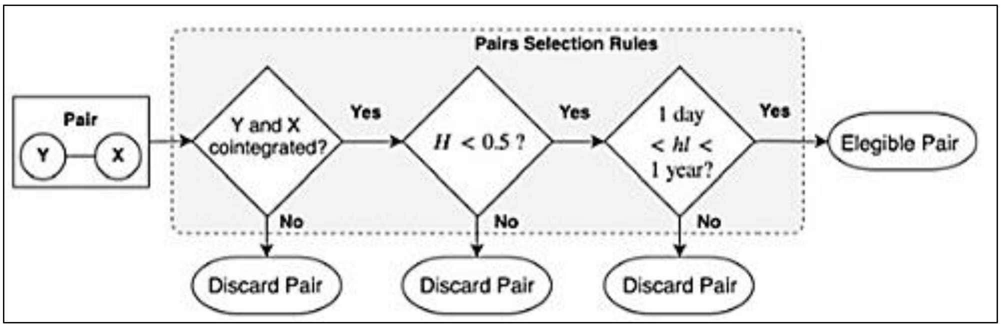


Fig 3.2.1 ARODs Process

Step 1 - The selection process starts with the testing of pairs, generated from the clustering step, for cointegration using the Engle-Granger test. This step reduced the number of pairs to 52.

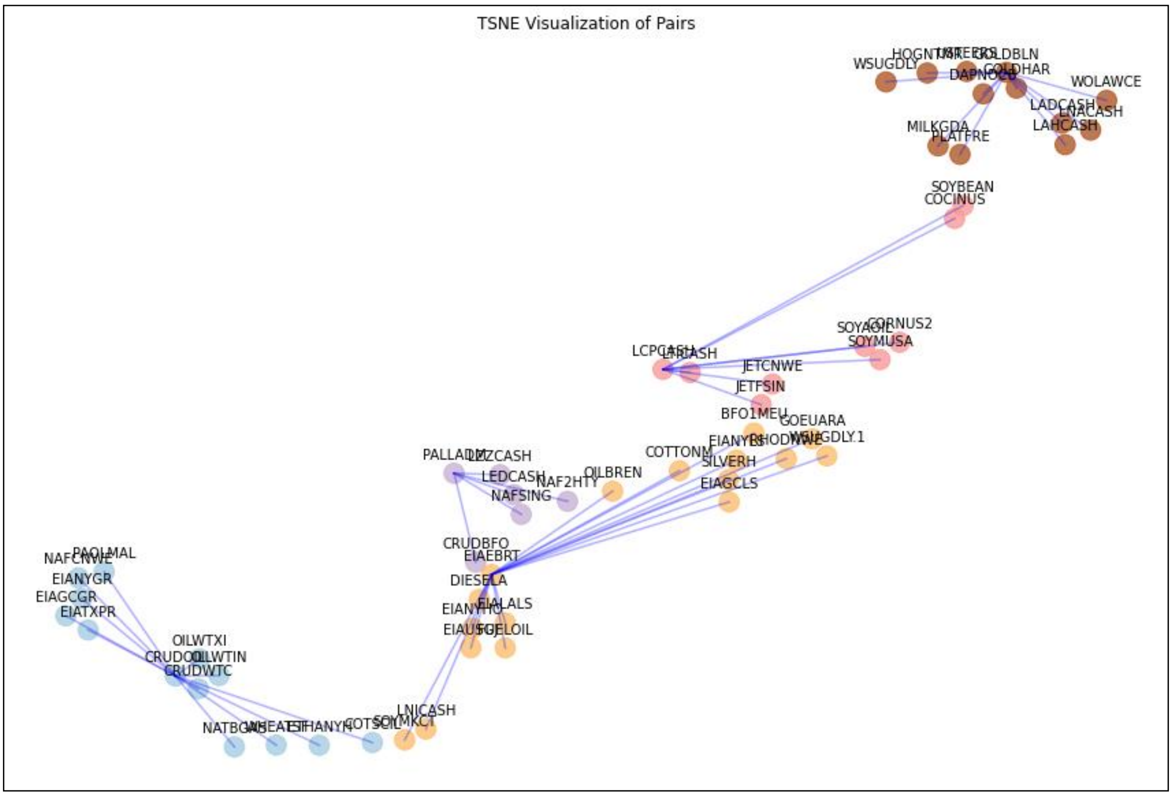


Fig 3.2.2 TSNE Visualization of 52 Pairs

Step 2 - A validation step is implemented to provide more confidence in the mean-reversion character of the pairs' spread. The condition imposed is that the Hurst exponent associated with the spread of a given pair is enforced to be smaller than 0.5, assuring the process leans towards mean-reversion. The Hurst exponent is a measure that quantifies the amount of memory in a specific time series (how mean-reverting is the process).

Step 3 - The pair's spread movement is constrained using the half-life of the mean-reverting process. The half-life is the time that the spread will take to mean-revert half of its distance after having diverged from the mean of the spread, given a historical window of data. For medium-term price movements, the spreads that have very short (< 7 days) or very long mean-reversion (> 365 days) periods are not suitable. The last 2 steps reduced the number of pairs to 27. The final pairs can be found in the Appendix.

4. Trading Models

4.1. Ordinary Least Squares (OLS)

The hedge ratio for the 27 pairs is derived by running OLS on the training data. It is static and will be used for backtesting the trading data. The trading signal will be the standard deviation calculated from the training data.

The trading strategy will long/short the spread if it is less/more than 0 standard deviation and stop-loss if the spread is more than 2 standard deviations. The stop loss is implemented because there could be structural breaks or regime changes that could make the static hedge ratio unstable. The timeframe for training is from 1/1/2017 to 31/12/2017, while the trading is from 1/1/2018 to 30/4/2022

From the backtesting, 2 pairs gave positive PnL.

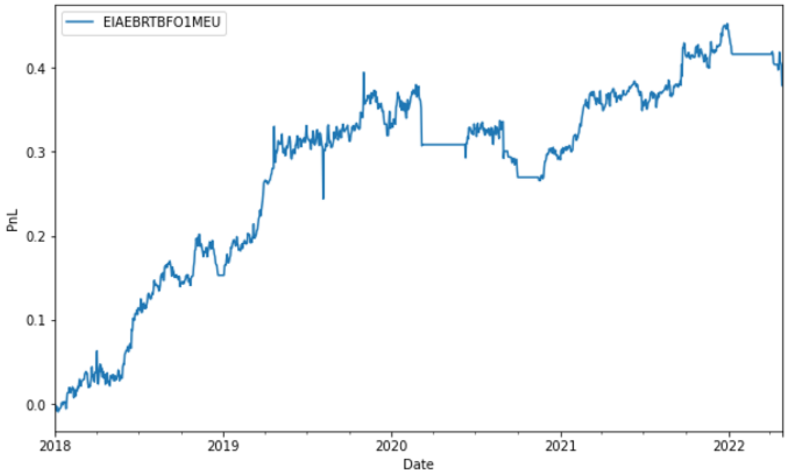


Fig 4.1.1 OLS Model – EIAEBRT-BFO1MEU (Cumulative Profit)

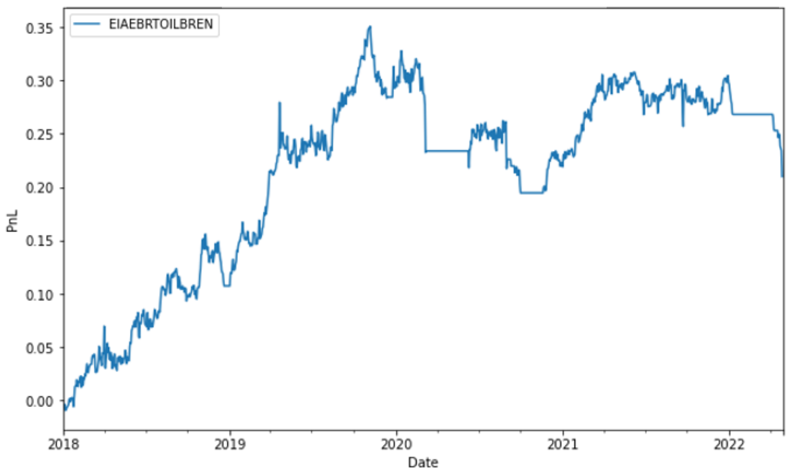


Fig 4.1.2 OLS Model – EIAEBRT-OILBREN (Cumulative Profit)

Pair Name	Cum. Return	Sharpe ratio	Sortino ratio	Max drawdown
EIAEBRT-BFO1MEU	0.378	0.755	0.068	0.126
EIAEBRT-OILBREN	0.209	0.468	0.040	0.149

Table 4.1.1 OLS Model – Performance Metrics

Though the results were not impressive (i.e. low cumulative return and Sharpe ratio), it is noteworthy that these two pairs have consistently been generating profit with low max drawdown. Hence, it is possible to leverage up the pairs to generate more returns, possibly between 5x to 10x. Furthermore, the underlying of both pairs are from the same industry, specifically in crude oil, implying that both pairs are more reliable to trade for fundamental reasoning.

All the remaining pairs either did not generate trade signals or were making losses. This is well within our expectation as the fundamental flaw of this method is to trade on a static hedge ratio.

4.2. Rolling Ordinary Least Squares with Bollinger Bands (R-OLS)

The R-OLS model is an enhanced version of the OLS model. Instead of using 1 set of training periods for trading, R-OLS applies OLS across fixed windows of observations and then rolls (moves or slides) the window across the data set. As a result, every trading day will have a unique hedge ratio value used for each pair. The R-OLS will be applied with Bollinger Bands to decide the entry and exit points.

The R-OLS model will follow trading parameters with the OLS model so that the relative effectiveness of the respective trading models can be compared.

The trading strategy will long/short the spread if it is less/more than 0 standard deviation and stop-loss if the spread is more than the rolling 2 standard deviations. The rolling window is set at 252 days while the timeframe for trading is from 1/1/2018 to 30/4/2022.

From the backtesting, 20 pairs gave positive PnL. For comparing purposes, only EIAEBRT-BFO1MEU and EIAEBRT-OILBREN pairs will be shown.

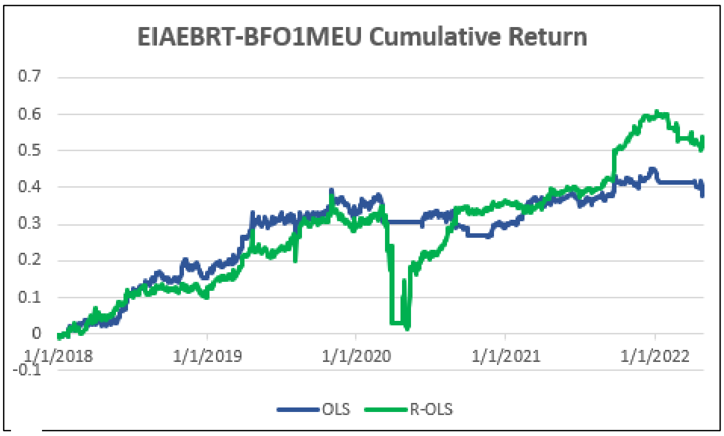


Fig 4.2.1 R-OLS Model – EIAEBRT-BFO1MEU (Cumulative Profit)

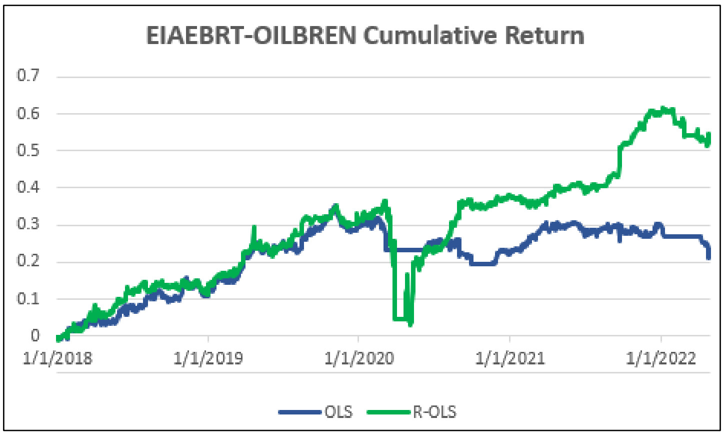


Fig 4.2.2 R-OLS Model – EIAEBRT-OILBREN (Cumulative Profit)

Pair Name	Cum. Return	Sharpe ratio	Sortino ratio	Max drawdown
EIAEBRT-BFO1MEU (OLS)	0.378	0.755	0.068	0.126
EIAEBRT-BFO1MEU (R-OLS)	0.537	0.674	0.053	0.333
EIAEBRT-OILBREN (OLS)	0.209	0.468	0.040	0.149
EIAEBRT-OILBREN (R-OLS)	0.548	0.716	0.056	0.313

Table 4.2.1 R-OLS Model – Performance Metrics

From the table above comparing the 2 different models, and the fact that R-OLS produced more pairs with positive PnL, it can be concluded that R-OLS is a better approach. It can adapt to any new changes in the relationship between the pairs and result in better performance.

4.3. Kalman Filter with Bollinger Band

The Kalman Filter is an optimal linear algorithm that updates the expected value of a hidden variable based on the latest value of an observable variable. The hidden variable, in this case, is the hedge ratio of each pair and the observable variables will be one of the components of the pair. Kalman filter works by minimizing the mean squared errors of estimated parameters. It is also a filter because it filters out noise from data to find the best estimate (Lee, 2021).

The reason behind the usage of the Kalman Filter is to reduce the arbitrariness of hedge ratio forecasting that happens with the R-OLS model. With Kalman Filter, there is no need to arbitrarily select the rolling window and the importance of different data points in the selected period (Wang, 2022).

For example, the scatterplot below shows the changes in the relationship between EIAEBRT-OILBREN pairs from 2009 to 2017.

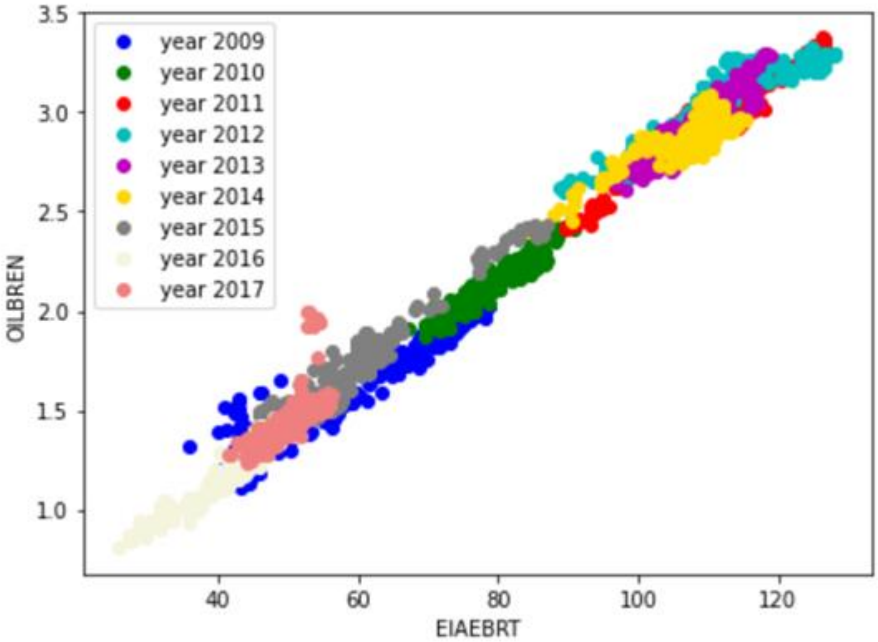


Fig 4.3.1 Kalman Filter – EIAEBRT-BFO1MEU (Relationship)

From the scatter plot, it can be observed that the relationship changes through time. Therefore, the hedge ratio changes over time and any pairs trading strategy should adapt to it.

The diagram below shows how the Kalman Filter slope and intercept throughout the years.

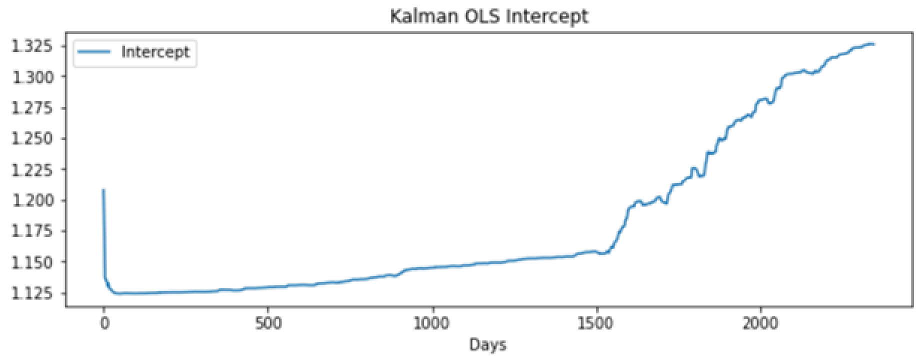


Fig 4.3.2 Kalman Filter – EIAEBRT-BFO1MEU (Intercept)

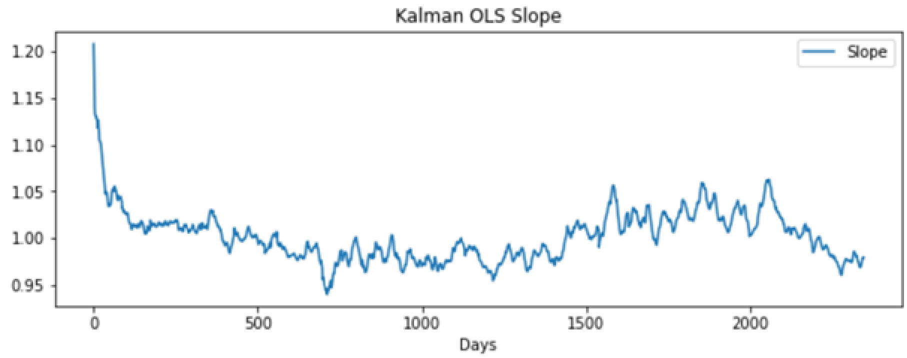


Fig 4.3.3 Kalman Filter – EIAEBRT-BFO1MEU (Slope)

From the diagram, it can be observed that both the intercept and slope are constantly changing. With Kalman Filter, it is possible to get the optimized hedge ratio as it updates the intercept and slope continuously.

The Kalman Filter will be applied with Bollinger Bands to decide the entry and exit points.

The trading strategy will long/short the spread if it is less/more than the rolling 1.5 standard deviation and stop-loss if the spread is more than 5% of allocated capital. The rolling window is set at 125 days while the timeframe for trading is from 1/1/2018 to 30/4/2022.

From the backtesting, 19 pairs gave positive PnL.

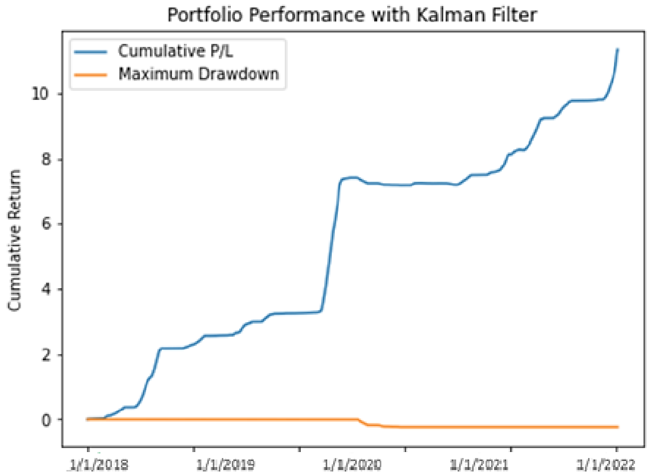


Fig 4.3.4 Kalman Filter Model – Equal Weighted Portfolio (Cumulative Return)

Pair Name	Cum. Return	Sharpe ratio	Sortino ratio	Max drawdown
Equal Weighted Portfolio	9.588	0.444	7.513	0.027

Table 4.3.1 Kalman Filter Model – Performance Metrics

Most profitable pairs came from oil and its derivative product. Some profitable unique pairs came from a different group of commodities such as LCPCASH-GOEUARA or LCPCASH-JETCNWE. One reason for this is copper is heavily used in oil and gas processing facilities.

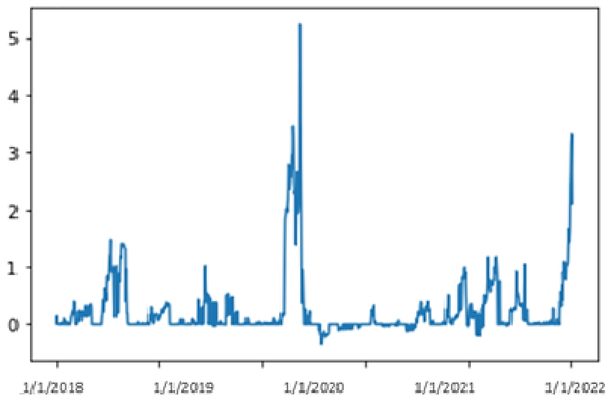


Fig 4.3.5 Kalman Filter Model – Equal Weighted Portfolio (Volatility)

Upon closer inspection of the portfolio volatility, it can be observed that huge volatility in the year 2020 and 2022 that benefited the strategy. The portfolio generated a high potential upside return during volatile times while only faced a few days of losses as seen from the maximum drawdown.

4.4. Support Vector Machine with Markov Regime Switching (SVM)

The idea behind this trading model is to generate dynamic hedge ratios using Support Vector Regression (SVR) for each period that is identified by the Markov Regression as a switch in the regime. After which, a new spread series for each pair using the hedge ratios is computed.

Markov-Switching Models

Regime changes can be described as a break (temporarily) between the historic relationship between two assets (i.e. changes in mean and variance), and their spread can deviate significantly from its historical equilibrium leading to a new equilibrium state. There may be a myriad of reasons as to why a temporary regime change occurs, but common reasons include financial crisis, period of economic recession or market crash (e.g. COVID-19 Pandemic).

Traditional pairs trading strategies are unable to detect such 'regime changes' which can lead to big losses. Using Markov regime-switching models allows the detection of such changes and adapts the trading rules accordingly. It should also be noted that regime change models differ from structural break models.

Structural Break Models vs. Regime Change Models

Regime Change Models	Structural Break Models
• Parameters vary across different regimes. • Finite number of regimes. • Regimes can be temporary and recurring. • Used to model effects of cyclically occurring changes in the economy. • Regime is unobserved and driven by stochastic process.	• Parameters change at different times. • Infinite number of structural changes. • Non-recurring and permanent shifts. • Usually used to model effects of permanent changes in economic structure.

Fig 4.4.1 Difference Between Regime Change & Structural Break Models

The Markov-switching model is a popular type of regime-switching model which assumes that unobserved states are determined by an underlying stochastic process known as a Markov-chain. A Markov-chain is a stochastic process used to describe how uncertain and unobserved outcomes occur. In the case of the Markov-switching model, it is used to describe how data falls into unobserved regimes.

A Markov-chain has the property that future states are dependent only on present states (this is known as the Markov property). A key characteristic of a Markov-chain is the transition probabilities. The transition probabilities describe the likelihood that the current regime stays the same or changes (i.e. the probability that the regime transitions to another regime).

Identifying Regime Changes

This section will run through the process of identifying regime changes. The pair EIAEBRT-BFO1MEU will be used as an example.

STEP 1: Construct the Spread Ratio

The spread is constructed as a price ratio of the two commodities in the pair (i.e. raw prices of EIABERT divided by raw prices of BFO1MEU, P_t^A/P_t^B). Other ways of constructing the spread series $(P_t^A/P_0^A) - \beta(P_t^B/P_0^B)$ or $\ln(P_t^A/P_0^A) - \beta \ln(P_t^B/P_0^B)$.

The plot of the ratio is as follows:

Fig 4.4.2 Markov Switching Model – EIAEBRT-BFO1MEU (Spread Ratio)

STEP 2: Fitting the Markov Switching Model

The Markov regression can be represented as:

$$y_t = a_{S_t} + x_t' \beta_{S_t} + \varepsilon_t$$
$$where\ \varepsilon_t \sim N\left(0, \sigma_{S_t}^2\right)$$

After feeding our spread ratio series into the model and setting the parameters of ‘K_Regimes’ as 2 (i.e. high and low mean regimes) and switching variance as TRUE, the following regression summary is created:

```
                        Markov Switching Model Results
==============================================================================
Dep. Variable:                      y   No. Observations:                 3586
Model:               MarkovRegression   Log Likelihood                8465.914
Date:                Sun, 19 Jun 2022   AIC                         -16919.828
Time:                        17:34:23   BIC                         -16882.720
Sample:                    08-01-2008   HQIC                        -16906.601
                         - 04-29-2022
Covariance Type:               approx
                             Regime 0 parameters
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          1.0012      0.000   2949.902      0.000       1.001       1.002
sigma2         0.0002   6.65e-06     23.652      0.000       0.000       0.000
                             Regime 1 parameters
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
const          0.9551      0.001    684.529      0.000       0.952       0.958
sigma2         0.0024   9.21e-05     26.135      0.000       0.002       0.003
                         Regime transition parameters
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
p[0->0]        0.9925      0.002    500.723      0.000       0.989       0.996
p[1->0]        0.0104      0.003      3.786      0.000       0.005       0.016
==============================================================================
```

Fig 4.4.3 Markov Switching Model – EIAEBRT-BFO1MEU (Regression Summary)

The coefficient of both the constants refers to the mean in the high and low regime, with the higher value being assigned to the high mean regime.

STEP 3: Plotting the Spread Ratio & Transition Probabilities

With the plots, it is easier to visualize the probability of high-mean regimes. Every time the spread ratio crosse the high mean coefficient of 1.0012, that period is considered the ‘high’ regime.

Fig 4.4.4 Markov Switching Model – EIAEBRT-BFO1MEU (Spread Ratio & Transition Probability)

Naturally, it would be ideal to isolate these dates where the regime switches from one state to another across all pairs. It should be noted that from 27 pairs from our pair selection process, only left with 15 pairs converged using the Markov Regression. The 12 pairs that did not converge were discarded.

EIAEBRT-OILBREN	CRUDOIL-NAFCNWE	LCPCASH-GOEUARA
EIAEBRT-BFO1MEU	LCPCASH-JETCNWE	WHEATSF-NAFCNWE
EIAEBRT-GOEUARA	LCPCASH-LTICASH	WHEATSF-EIANYGR
EIAEBRT-EIALALS	PALLADM-LZZCASH	WHEATSF-EIAGCGR
EIAEBRT-DIESELA	LNICASH-RHODNWE	WHEATSF-NATBGAS

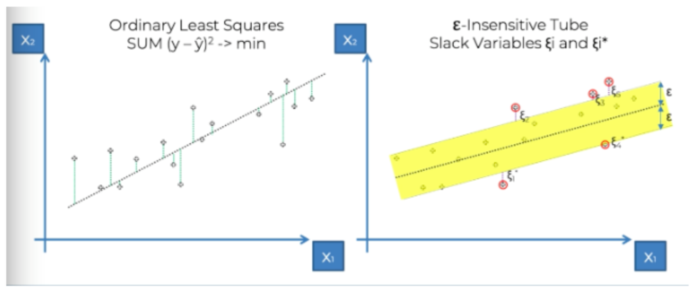
Table 4.4.1 Markov Regression Trading Pairs

	EIAEBRT-OILBREN	EIAEBRT-BFO1MEU	EIAEBRT-GOEUARA	EIAEBRT-EIALALS	EIAEBRT-DIESELA	CRUDOIL-NAFCNWE	LCPCASH-JETCNWE	LCPCASH-LTICASH
0	2008-08-01	2008-08-01	2008-08-01	2008-08-01	2008-08-01	2008-08-01	2008-08-01	2008-08-01
1	2008-09-01	2010-06-28	2009-02-27	2020-03-31	2020-03-31	2009-01-12	2015-12-09	2008-10-23
2	2008-09-02	2010-08-19	2009-03-19	2020-04-01	2020-04-01	2009-01-23	2016-04-29	2009-03-13
3	2008-10-15	2010-08-27	2009-04-06	2020-04-21	2020-04-21	2009-01-27	2016-11-08	2009-05-07

Fig 4.4.5 Markov Switching Model – Regime Periods (Not Full List)

Support Vector Machine

For every single regime period, SVM is used to generate the respective hedge ratio. SVM is chosen because it demonstrates superior performance as compared to OLS (Baek, Glambosky, Oh and Lee, 2020). While linear regression models minimize the error between the actual and predicted values through the line of best fit, SVM manages to fit the best line within a threshold of values, otherwise called the epsilon-insensitive tube.



Ordinary Least Squares. vs Epsilon-Insensitive Tube

Fig 4.4.6 Difference Between SVM & OLS

	EIAEBRT-OILBREN	EIAEBRT-BFO1MEU	EIAEBRT-GOEUARA	EIAEBRT-EIALALS	EIAEBRT-DIESELA	CRUDOIL-NAFCNWE	LCPCASH-JETCNWE	LCPCASH-LTICASH	PALLADM-LZZCASH	LNICASH-RHODNWE	LCPCASH-GOEUARA	WHEATSF-NAFCNWE	WHEATSF-EIANYGR
0	1.001373	1.038570	0.090530	39.964544	40.233918	0.177837	7.047187	0.359216	0.234107	3.277812	6.656740	0.006820	1.850302
1	0.994594	0.998600	0.118954	2.062000	2.061500	0.128090	13.381932	0.285507	0.302627	11.208335	6.009769	0.011339	1.499369
2	0.986521	0.978831	0.107602	9.564152	9.545173	0.115760	10.630528	0.377249	NaN	NaN	6.695555	0.006324	1.290831
3	0.984428	0.987244	0.118348	1.297000	1.296800	0.106326	11.834905	0.329900	NaN	NaN	9.800936	0.012341	1.251356
4	1.027258	0.973257	0.114955	NaN	NaN	0.130875	11.395473	0.408744	NaN	NaN	8.567690	0.010139	2.020676

Fig 4.4.5 Markov Switching Model – Hedge Ratio

The spread for each pair across the entire sample period is based on the generated hedge ratios:

$$Spread = log(a) - nlog(b)$$

where 'a' and 'b' are prices of futures A and B respectively and n is the hedge ratio

Date	EIAEBRT-OILBREN	EIAEBRT-BFO1MEU	EIAEBRT-GOEUARA	EIAEBRT-EIALALS	EIAEBRT-DIESELA	CRUDOIL-NAFCNWE	LCPCASH-JETCNWE	LCPCASH-LTICASH	PALLADM-LZZCASH	LNICASH-RHODNWE	LCPCASH-GOEUARA	WHEATSF-NAFCNWE
2008-08-01	0.042368	-0.164151	4.140996	-45.303462	-44.688611	3.723899	-44.805737	5.318034	4.724704	-11.195798	-42.436995	1.722251
2008-08-04	0.042726	-0.164838	4.122488	-44.232787	-43.913918	3.716466	-44.860795	5.292291	4.712232	-11.146539	-42.471703	1.669710
2008-08-05	0.037160	-0.167783	4.083050	-43.037082	-43.037133	3.699428	-44.410690	5.305565	4.697474	-11.140387	-42.042167	1.737918
2008-08-06	0.038941	-0.165821	4.066628	-42.407230	-42.362590	3.694465	-44.329836	5.297896	4.723640	-11.098889	-41.954926	1.722598
2008-08-07	0.039912	-0.161441	4.087419	-42.410504	-41.807905	3.708620	-44.229274	5.302608	4.698953	-11.040070	-41.991473	1.785820
...	...	...	...	...	...	...	...	...	...	...	...	...
2022-04-25	-0.296024	-0.121124	3.725034	2.841332	2.818249	3.802615	-100.211263	5.129523	5.514216	-80.993133	-74.198683	2.306973
2022-04-26	-0.289393	-0.113449	3.754334	2.841161	2.799756	3.821579	-100.975585	5.130941	5.546098	-80.877721	-74.813157	2.326149
2022-04-27	-0.287929	-0.111896	3.756656	2.809804	2.769630	3.820943	-101.239925	5.135131	5.574531	-80.870907	-74.971017	2.324011
2022-04-28	-0.289015	-0.112095	3.777924	2.810599	2.771127	3.853157	-101.485958	5.117132	5.577217	-80.881434	-75.223039	2.319414
2022-04-29	-0.264917	-0.087997	3.800647	2.836878	2.797309	3.844324	-101.726184	5.123570	5.607657	-80.918883	-75.346696	2.294016

Fig 4.4.6 Markov Switching Model – Pair Spread

Backtesting

The trading strategy will long/short the spread if the rolling mean it is less/more than the spread rolling standard deviation. The rolling window is set at 42 days while the timeframe for trading is from 1/1/2008 to 30/4/2022.

From the backtesting, 11 pairs gave positive PnL.

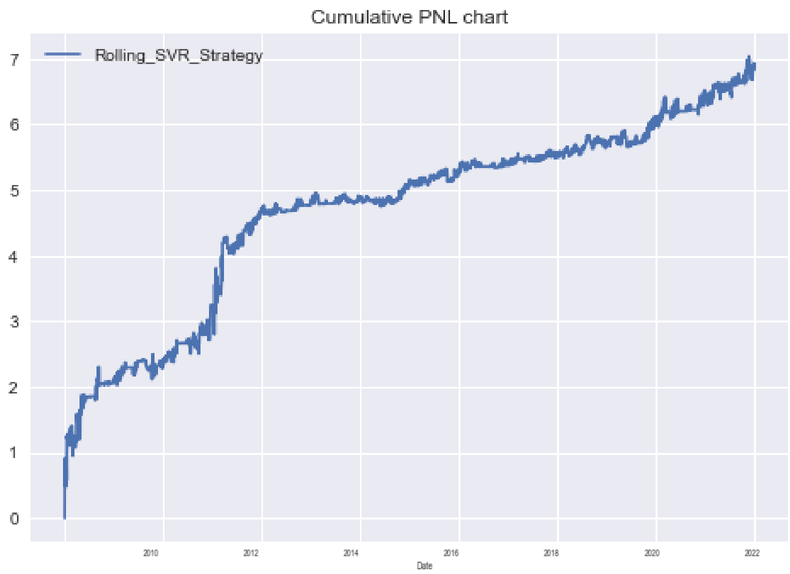


Fig 4.4.7 Markov Switching Model – Equal Weighted Portfolio (Cumulative Return)

Pair Name	Cum. Return	Sharpe ratio	Sortino ratio	Max drawdown
Equal Weighted Portfolio	6.850	0.583313	-	0.519068

Table 4.3.1 Kalman Filter Model – Performance Metrics

5. Reference

Harlacher, M. (2016). Cointegration based algorithmic pairs trading. PhD thesis, University of St. Gallen.

Sarmento, S. M., & Horta, N. (2020). Enhancing a pairs trading strategy with the application of machine learning. Expert Systems with Applications, 158, 113490. https://doi.org/10.1016/j.eswa.2020.113490

Baek, Seungho, Mina Glambosky, Seok H. Oh, and Jeong Lee. (2020). Machine Learning and Algorithmic Pairs Trading in Futures Markets. Sustainability 12, no. 17: 6791. https://doi.org/10.3390/su12176791

Lee, O. (2021). Kalman Filters In Pairs Trading. Medium. Retrieved 25 June 2022, from https://medium.datadriveninvestor.com/kalman-filters-in-pairs-trading-dcf98a91a808.

Wang, L. (2022). Kalman Filter and Pairs Trading. Quantitative Trading and Systematic Investing. Retrieved 25 June 2022, from https://letianzj.github.io/kalman-filter-pairs-trading.html.

6. Appendix

List of Commodity Futures

No	Name	Type	Sub-Type	Ticker
1	Steel Iron ore Fe62% AUS CIF China	Metal	Steel	SHCNI62
2	Steel Iron ore Fe65% BR CIF China	Metal	Steel	SHCNI58
3	LME-Copper Grade A Cash U$/MT	Metal	Non-Ferrous Metal	LCPCASH
4	LME-Nickel Cash U$/MT	Metal	Non-Ferrous Metal	LNICASH
5	LME-NASAAC Cash U$/MT	Metal	Non-Ferrous Metal	LNACASH
6	LME-Aluminium 99.7% Cash U$/MT	Metal	Non-Ferrous Metal	LAHCASH
7	LME-Aluminium Alloy Cash U$/MT	Metal	Non-Ferrous Metal	LADCASH
8	LME-SHG Zinc 99.995% Cash U$/MT	Metal	Non-Ferrous Metal	LZZCASH
9	LME-Tin 99.85% Cash U$/MT	Metal	Non-Ferrous Metal	LTICASH
10	LME-Lead Cash U$/MT	Metal	Non-Ferrous Metal	LEDCASH
11	Rhodium CIF NWE U$/Ounce	Metal	Precious Metals	RHODNWE
12	Gold Bullion LBM $/t oz DELAY	Metal	Precious Metals	GOLDBLN
13	Silver, Handy&Harman (NY) U$/Troy OZ	Metal	Precious Metals	SILVERH
14	Palladium U$/Troy Ounce	Metal	Precious Metals	PALLADM
15	London Platinum Free Market $/Troy oz	Metal	Precious Metals	PLATFRE
16	Gold, Handy & Harman Base $/Troy Oz	Metal	Precious Metals	GOLDHAR
17	Ethanol, Spot NY Harbour U$/GAL	Chemical	Ethanol	ETHANYH
18	DAP, New Orleans CFR Barge U$/MT	Chemical	Fertilizer	DAPNOCB
19	Urea Granular CFR New Orleans $/MT	Chemical	Fertilizer	UREAGRN
20	USGC GLN REG Spt Price FOB U$/GAL	Energy	Gas	EIAGCGR
21	NY Conv GLN REG Spt Price FOB U$/GAL	Energy	Gas	EIANYGR
22	ICE Natural Gas 1 Mth.Fwd. P/Therm	Energy	Gas	NATBGAS
23	Crude Oil WTI NYMEX Close M U$/BBL	Energy	Crude Oil	OILWTXI
24	Crude Oil North Sea BFO FOB U$/BBL	Energy	Crude Oil	CRUDBFO
25	Crude Oil BFO M1 Europe FOB $/Bbl	Energy	Crude Oil	BFO1MEU
26	Crude Oil WTI FOB Cushing U$/BBL	Energy	Crude Oil	CRUDWTC
27	Crude Oil-WTI Spot Cushing U$/BBL	Energy	Crude Oil	CRUDOIL
28	Europe Brent Spot FOB U$/BBL Daily	Energy	Crude Oil	EIAEBRT
29	Crude Oil BFO M1 Europe FOB $/BBl	Energy	Crude Oil	OILBREN
30	Gasoil, 0.2% Sulphur FOB ARA U$/MT	Energy	Fuel Oil	GOEUARA
31	Fuel Oil No.2 (New York) C/Gallon	Energy	Fuel Oil	FUELOIL
32	NY No. 2 HO Spt Price FOB U$/GAL	Energy	Heating Oil	EIANYHO
33	Diesel, .05% Sulphur LA C/GAL	Energy	Diesel	DIESELA
34	SNL US Electricity Peak load SP-15	Energy	Electricity	ES15PSN
35	EEX - Phelix Peak Hr.09-20 E/Mwh	Energy	Electricity	EEXPEAK
36	Electricity PJM Base Rate U$/MWh	Energy	Electricity	ELEPJMB
37	Electricity PJM Peak Rate U$/MWh	Energy	Electricity	ELEPJMP
38	EEX - Phelix Base Hr.01-24 E/Mwh	Energy	Electricity	EEXBASE
39	USGC KERO Jet Spt Price FOB U$/GAL	Energy	Jetfuel	EIAUSGJ
40	Jet Kerosene CIF NWE U$/MT	Energy	Jetfuel	JETCNWE
41	Jet Kerosene FOB Singapore U$/BBL	Energy	Jetfuel	JETFSIN
42	LA ULSD CARB Spot Price U$/GAL	Energy	Sulphur	EIALALS

43	NY ULSD No. 2 Spot Price U$/GAL	Energy	Sulphur	EIANYLS
44	USGC ULSD No. 2 Spot Price U$/GAL	Energy	Sulphur	EIAGCLS
45	Naphtha 2 Half Tokyo Near M C+F $/MT	Energy	Naphtha	NAF2HTY
46	Naphtha, FOB Singapore U$/BBL	Energy	Naphtha	NAFSING
47	Naphtha, CIF NWE U$/MT	Energy	Naphtha	NAFCNWE
48	Mont Belvieu TX Prop Spt FOB U$/GAL	Energy	Propane	EIATXPR
49	Yellow Soybn US NO.1 Sth Dvprt U$/Bsh	Agriculture	Soy	SOYADSC
50	Soya Oil, Crude Decatur US $/lb	Agriculture	Soy	SOYAOIL
51	Soyameal USA 48% Protein $/MT	Agriculture	Soy	SOYMUSA
52	Soyabeans, No.1 Yellow $/Bushel	Agriculture	Soy	SOYBEAN
53	Soymeal 48% FOB K.City $/MT	Agriculture	Soy	SOYMKCT
54	Corn No.2 Yellow U$/Bushel	Agriculture	Corn	CORNUS2
55	Corn US No.2 South Central IL $/BSH	Agriculture	Corn	COTSCIL
56	Wheat US HRS 14% Del Mineapolis/Dulut	Agriculture	Wheat	WHTHRMD
57	Wheat No.2,Soft Red U$/Bu	Agriculture	Wheat	WHEATSF
58	Rice, White 100% FOB Bangkok U$/MT	Agriculture	Rice	WSUGDLY
59	Milk Non Fat Dry Grade A Spot	Agriculture	Milk	MILKGDA
60	Live Steers, USDA 5 Area Wtd. Avge.	Agriculture	Steers	USTEERS
61	HOG 51-52% US 3 AREA Ntnl MR U$/Cwt	Agriculture	Hog	HOGNTMR
62	Wool AWEX E.M.I. A$/100KG	Agriculture	Wool	WOLAWCE
63	Palm Kernel Oil MAL CIF Rdam US$ /MT	Agriculture	Palm Oil	PAOLMAL
64	Raw Sugar-ISA Daily Price c/lb	Agriculture	Sugar	WSUGDLY
65	Cocoa-ICCO Daily Price US$/MT	Agriculture	Cocoa	COCINUS
66	Cotton,1 1/16Str Low -Midl,Memph $/Lb	Agriculture	Cotton	COTTONM

Table A.1 List of Commodity Futures

List of Final Trading Pairs

No	Leg1	Leg2	Trading Pairs
1	EIAEBRT	OILBREN	EIAEBRT-OILBREN
2	EIAEBRT	EIAUSGJ	EIAEBRT-EIAUSGJ
3	EIAEBRT	BFO1MEU	EIAEBRT-BFO1MEU
4	EIAEBRT	GOEUARA	EIAEBRT-GOEUARA
5	EIAEBRT	EIALALS	EIAEBRT-EIALALS
6	EIAEBRT	EIANYHO	EIAEBRT-EIANYHO
7	EIAEBRT	EIANYLS	EIAEBRT-EIANYLS
8	EIAEBRT	EIAGCLS	EIAEBRT-EIAGCLS
9	EIAEBRT	DIESELA	EIAEBRT-DIESELA
10	EIAEBRT	FUELOIL	EIAEBRT-FUELOIL
11	CRUDOIL	OILWTIN	CRUDOIL-OILWTIN
12	CRUDOIL	NAFCNWE	CRUDOIL-NAFCNWE
13	CRUDOIL	CRUDWTC	CRUDOIL-CRUDWTC
14	CRUDOIL	ETHANYH	CRUDOIL-ETHANYH
15	CRUDOIL	EIANYGR	CRUDOIL-EIANYGR
16	CRUDOIL	EIAGCGR	CRUDOIL-EIAGCGR
17	CRUDOIL	OILWTXI	CRUDOIL-OILWTXI
18	GOLDBLN	GOLDHAR	GOLDBLN-GOLDHAR
19	LCPCASH	JETCNWE	LCPCASH-JETCNWE
20	LCPCASH	LTICASH	LCPCASH-LTICASH
21	PALLADM	LZZCASH	PALLADM-LZZCASH
22	LNICASH	RHODNWE	LNICASH-RHODNWE
23	LCPCASH	GOEUARA	LCPCASH-GOEUARA
24	WHEATSF	NAFCNWE	WHEATSF-NAFCNWE
25	WHEATSF	EIANYGR	WHEATSF-EIANYGR
26	WHEATSF	EIAGCGR	WHEATSF-EIAGCGR
27	WHEATSF	NATBGAS	WHEATSF-NATBGAS

Table A.2 List of Final Trading Pairs