Final Project

Mean Variance Optimization in Commodities Rolling and Optimized Portfolio Weight Selection

Jonah Pandarinath (30130045)

Haskayne School of Business, University of Calgary

FNCE 449: Trading and Data Management

Prof. S George Hart

November 3rd, 2023

Table of Contents

Overview

Portfolio optimization is a rather niche topic within finance. It is the intersection of statistics, data science, financial theory, econometrics, and mathematics. Throughout, I explore the two topics of contract roll optimization for commodities, and how the impact of a static roll strategy has when compared to a portfolio risk standpoint. This project also stands as a starting point towards creating a risk parity portfolio, which optimizes asset class weights, as well as individual security weights within those asset classes.

Commodity roll optimization analyzes the impact of a single day roll, compared against multi day roll strategies over a seven-year period. The goal is to identify a static strategy for what is the best day to roll a contract of a specific commodity. The hypothesis is that each commodity will have different 'optimized' roll dates which can be exploited to maximise the mean-variance of rolling a specific commodity's contracts.

The asset allocation and covariance shrinkage analysis centres around a single asset class of commodity futures contracts, and how to optimize portfolio weights a modified (shrunk) version of the covariance matrix. This shrunken covariance matrix exploits specific properties within financial market return data to reduce estimation error in the covariance matrix. This exercise also stands as a starting point towards an eventual exploration of a multi-asset class risk parity portfolio.

Investible Universe:

Energy	Metals	Grains and Oilseeds	Other
Crude Oil WTI (CL Comdty)	Silver (SI Comdty)	Canola (RS Comdty)	Sugar (SB Comdty)
Natural Gas (NG Comdty)	Platinum (PL Comdty)	Corn (C Comdty)	Lean Hogs (LH Comdty)
Heating Oil (HO Comdty)	Copper (HG Comdty)	Soybeans (S Comdty)	Cotton (CT Comdty)
Ethanol (CUA Comdty)	Palladium (PA Comdty)	Wheat (W Comdty)	VIX (UX Comdty)
	Gold (GC Comdty)	Oats (O Comdty)	Coffee (KC Comdty)
		Rough Rice (RR Comdty)	Feeder Cattle (FC Comdty)

Data Collection and Testing

Due to data availability, a period of 10-years of historical price returns of commodity futures contracts is analyzed using data obtained from Bloomberg. The data sample is split for all optimization calculations into a training period from September 2013 to September 2020, and a testing period from September 2020 to September 2023. This ensures that the optimization is not overfit to the data and assumes live practice where a trader is unaware of future price returns for an underlying. Additionally, it allows for algorithm testing, wherein an algorithm will trade following a set of rules trained on previous data.

Data for the covariance matrix shrinkage is the calculated optimized contract rolling rules from the commodity rolling analysis through the testing period. The covariance calculated weights of each individual assets are by design not optimized with future unknown returns. The covariance matrices are calculated on a trailing p number of returns to predict the optimal weights for the next, or next few periods.

The entire project has been coded in Python for scalability and the repo can be found at the following link: https://github.com/Jonah-Panda/Commodity_Rolling_Optimization

Commodity Rolling Optimization

Assumptions:

- Fractional Contracts
- No transaction costs
- Cannot hold a contract till expiry

After having collected commodity futures contracts for the past 10 years, the next challenge was getting a continuous price for each commodity. This was solved using a Python script that creates a continuous price history based on either a single day rolling of a contract, or a multi-day rolling of contracts. The hypothesis is that each individual commodity will have an optimal roll day (or days), and to find the day (or days) that maximizes expected return, while minimizing volatility of the commodity.

Mean variance optimization was calculated using trading day returns and the following formula:

$$Maximum \left\{ \frac{Mean\ of\ Returns}{Variance\ of\ Returns} \right\}$$

The risk-free rate is omitted in this calculation as commodities trade on an international market and the underlying assets are not subject to an individual country's overnight rate. Additionally, including the risk-free rate in the mean variance calculation has a marginal change on the outcome since the testing period for the data is from 2013 to 2020, which was an environment which has had historically low risk-free rates. Lastly, since each commodity is being compared to itself with a slightly different transaction pattern, all price returns are highly correlated further diminishing significance of the risk-free rate.

All calculations were separated between a training period from September 2013 to September 2020, and a testing period from September 2020 to September 2023. This ensures that the optimization is not overfit to the data, and assumes live practice where a trader is unaware of future price returns for an underlying.

Roll Results

The commodities all optimized quite differently. The trends are shown below:

Commodity	Training mVar	Roll Days	Last Roll Date	Trend	Testing mVar
CL	0.053630271	25	25	Long Roll Early	-0.0534807
NG	0.065290933	10	16	Medium Roll Middle	0.01450901
SI	0.027539641	1	18	Short Roll Middle	-0.01187044
PL	0.099139498	1	9	Short Roll Early	0.00739965
SB	0.09818677	23	25	Long Roll Early	-0.28442389
HG	0.052106025	1	23	Short Roll Early	-0.08183398
RR	0.026189053	25	25	Long Roll Early	-0.28161483
LH	0.084752462	18	1	Long Roll Late	-0.16477203
PA	-0.071153185	1	16	Short Roll Middle	0.07907236
HO	0.10207048	25	25	Long Roll Early	0
CUA	0.096965606	8	1	Medium Roll Late	-0.18749964
GC	-0.142919006	3	25	Short Roll Early	-0.03686652
CT	0.137209896	25	25	Long Roll Early	-0.0364892

FC	0.097151427	25	1	Long Roll Late	-0.40463187
KC	0.045699732	2	1	Short Roll Late	-0.07403891
RS	0.202524704	25	23	Long Roll Early	-0.10021225
C	0.127519708	7	12	Medium Roll Middle	-0.15527964
S	0.122969137	8	15	Medium Roll Middle	-0.19058947
W	0.072773785	25	25	Long Roll Early	-0.00965263
O	0.046676876	25	25	Long Roll Early	-0.00022141
UX	0.031147518	1	2	Short Roll Late	0.05644168

There is a significant discrepancy between the trained model for optimization and the testing period. This suggests that a static mean-variance optimized model used is not well suited for predicting when is optimal to roll a series of contracts on an underlying commodity. Heatmaps for training and testing period mean-variance are compared in Appendix A – Commodity MVO Plots. Taylor (2016) analyzed an optimization strategy that is dynamic, and concludes that a static strategy with unchanging roll dates will underperform a dynamic strategy where a trader would watch the mean and variance of execution quality as days in the month tick forward and estimate when execution quality would peak and roll their position over 5-8 days centered around the predicted maximum execution quality day. This requires that the trader predicts the maximum execution quality day three to four days in advance which can be challenging due to the nature of out-of-sample forecasting.

The optimization sample size for the metals Silver, Platinum, Palladium, and Gold were significantly shorter than the rest of the commodities. This was due to liquidity constraints as well as a shorter contract overlap period. Further discussion can be found in Appendix B – Contract Liquidity Constraints.

Goldman and UBS Roll

The Goldman roll is a multi day futures contract roll strategy that assumes uniform weights of rolling between the fifth and ninth day till proceeding the expiry of the contract. The 5-day roll strategy has the goal of reducing single day roll volatility risk by spreading the transaction over many days, while minimizing transaction costs. In the backtest, the Goldman Roll underperformed many different multi-day strategies since the optimization patterns were different for every commodity, and there should not be a catch-all rule for when to roll all futures contracts. A visual representation compared to optimized roll days is highlighted in Appendix A – Commodity MVO Plots.

Similarly, the UBS roll strategy (uniform weights between the sixth and tenth DTE) also underperforms during backtesting and testing periods. The failure of both strategies in a backtest further suggests proves difficulty to fit a static roll strategy to a dynamic security such as a commodity.

Dynamic Commodity Rolling Strategies

With the results of the tests underperforming, this begs the question, is a dynamic roll strategy better? If so, how would a trader use more recent information to roll a position forward. One strategy would be to use a rolling one-year trailing optimization to predict the thirteenth month's optimized roll. Essentially this would look like a simple moving average for optimized trade execution days. Or potentially an exponential moving average of 12 months backwards, as well as a 3-year trail rolling trend. Or potentially, using past information is irrelevant in commodity markets since news and forward-looking information have a larger material impact on commodities compared to equity and fixed income markets.

Portfolio Optimization and Covariance Shrinkage

The Efficiency Frontier

The covariance matrix between assets is often the starting point for any portfolio optimization strategy. The main idea is to maximize an investors Sharpe ratio and fall on the efficiency frontier. However, there is often a disconnect between how an optimized efficiency frontier portfolio performs overtime as the returns and variance of individual assets shift. Diversification among different assets has the benefit of reducing any individual assets idiosyncratic risk and when combined with other assets, reduces the volatility of the portfolio. Under the hood of this combination lives the covariance matrix between all assets of the portfolio.

Trailing Datapoints

Current financial covariance matrix estimation literature has set a limit to the number of assets to sample size of 1 to 1 Ledoit and Wolf (2020). For example, a portfolio of 21 assets, should have at least 21 trailing datapoints for returns to calculate the covariance matrix. This means that there must be 21+1 historical prices for the assets preceding the date for the covariance calculation since one datapoint is lost at $t = 0$ in the calculation of the return from $t = 0$ to $t = 1$. The ratio of N assets to trailing datapoints is represented as a concentration ratio Large, Reeves, and Haji (n.d.) and can be written as $c = \frac{p}{N}$ where p represents the sample size of historical return datapoints. Note: reducing $c \leq 1$ will result in a covariance matrix that is non-invertible Large, Reeves, and Haji (n.d.), which is a requirement to calculate the minimum variance weights of the portfolio.

The frequency of trailing datapoints also affects the calculation of the covariance matrix. This was demonstrated in the exploration of the beta of thinly traded securities, which presented material differences between the beta estimation at the minute, daily, weekly, and monthly timeframes. A portion of the differences are explained through serial autocorrelation, where the assets price (and thus return) at a time $t = 1$ are dependent of the assets price (and thus return) at a time $t = t - 1$. Using a longer timeframe between each datapoint for all trailing datapoints reduces serial autocorrelation; however, this comes at the expense of underestimating momentum factors in a security.

Covariance Shrinkage

A sample covariance matrix to estimate the population covariance matrix is not perfect. Specifically, when the sample size relative to the number of assets approaches $c = 1$ the estimation error in the covariance matrix increases. Additionally, financial markets have a unique structure in which a security's returns and volatility are rarely stable over periods of time. Financial markets also have unique properties which can be exploited to reduce estimation errors in the sample covariance matrix. Specifically, Ledoit and Wolf (2020) state that stock returns have a factor-model structure, have an average positive correlation, and an average positive covariance. Using these properties Ledoit and Wolf have developed several covariance shrinkage methods to improve the accuracy of estimated covariance matrices.

The process of covariance shrinkage utilizes the sample covariance matrix as a base to locate the sample eigenvalues of the covariance matrix. A shrinkage estimator then scales the eigenvalues of the covariance matrix towards a center of eigenvalues present in the sample covariance matrix. The shrinkage of eigenvalues is necessary since "most estimators tend to overestimate the risk of low volatility, low correlation assets, while underestimating the risk of high volatility, high correlation assets." Large, Reeves, and Haji (n.d., Overview para 3). The new shrunken eigenvalues are placed on the diagonal of the matrix J and multiplied between the original eigenvector matrix S and the inverse eigenvector matrix S^{-1}. This effectively transforms:

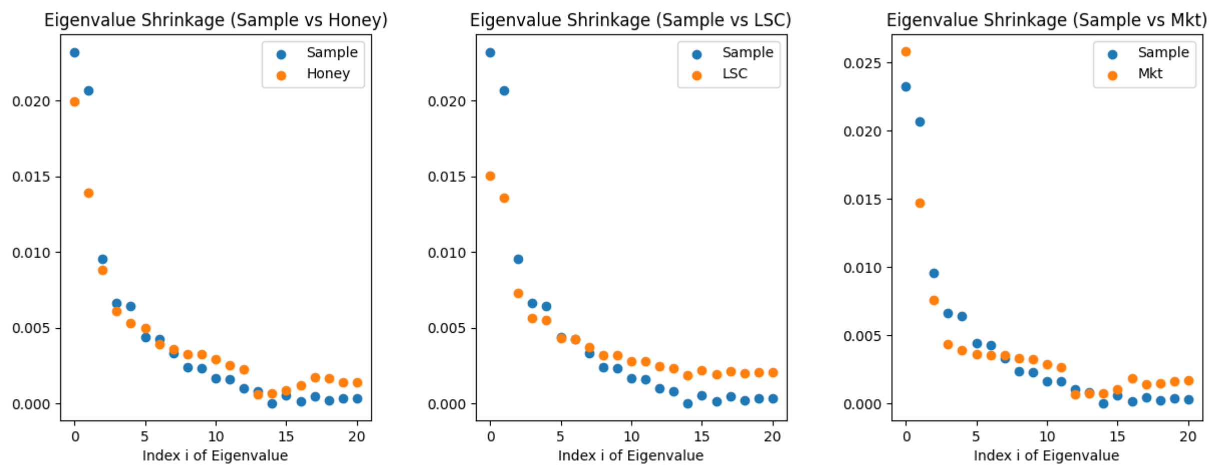
$$M = S\ E\ S^{-1}$$

Where M is the sample covariance matrix and E is the matrix with eigenvalues on the diagonal, and S is the matrix with the corresponding eigenvectors to the diagonal in E. To a more centered matrix with:

$$M_{Shrunk} = S\ J\ S^{-1}$$

A visual representation of eigenvalue shrinkage is presented in Figure 1 below, which highlights the eigenvalues of the original sample covariance matrix, as well as three linear forms of eigenvalue shrinkage. Recently, improvements in the field of random matrix theory has allowed for the development of non-linear eigenvalue shrinkage which is beyond the scope of this paper.

Figure 1 – Eigenvalue Shrinkage



Ledoit and Wolf (2004a) Ledoit and Wolf (2004b) Ledoit and Wolf (2003)

The "Honey" eigenvalue shrinkage is linear towards the mean constant correlation, Ledoit and Wolf (2004a). The "LSC" which stands for Large-Scaled Covariance shrinkage represents a linear shrinkage towards one-parameter matrix. Specific usage of this linear shrinkage model is for large-dimension covariance matrices (of which a 21-asset portfolio is not), Ledoit and Wolf (2004b). The "Mkt" eigenvalue shrinkage is a linear shrinkage towards a one-factor market model where the market is calculated as the mean of returns of all assets for each period, Ledoit and Wolf (2003). The market shrinkage is better suited for larger portfolios and is based on the CAPM. All further analysis will show all three-methods of shrinkage, however the best suited shrinkage method in this case is the "Honey" eigenvalue shrinkage since all the assets are of the same asset class.

Covariance Matrix Weights

The Global Minimum Variance (GMV) portfolio falls on the efficiency frontier, albeit not at the optimal point when it comes to a Sharpe ratio perspective. However, the benefit of calculating the GMV removes the requirement to estimate the expected return. Estimating the expected return is subject to errors in estimation, as well as the coined financial disclaimer phrase “past performance is not indicative of future performance”. The GMV asset weights are calculated as follows:

$$\text{Let } B^{-1} = {M_{Shrunk}}^{-1}$$

$$w_i = \frac{\sum row(B^{-1})}{\sum B^{-1}}$$

w_i represents the weight of each security given a total effect of its volatility and covariance with all other assets in the portfolio, Large, Reeves, and Haji (n.d.).

Weight of Weights

Selecting a timeframe for trailing datapoints has a material impact on the weight selection of a security. Therefore, it may be better to create a blend between different weights calculated from different shrunk covariance matrices. A covariance matrix using a concentration ratio $c = \frac{42}{21}$ has 42 sample datapoints. At the daily frequency, with two week rebalances, this covariance matrix would need to estimate the weights of a minimum volatility portfolio up to 10 periods in advance. Additionally, this represents an overlap of $\frac{10}{42} \approx 24\%$ of the sample size. Contrastingly, a covariance matrix with the same concentration ratio, with two-week frequency on sample returns data must only optimize the weights 1 period in advance before the next rebalance and has a prediction/sample overlap of $\frac{1}{42} \approx 2.4\%$. I suggest a balance formula between the two timeframes predicted weights as follows:

$$w_i = w_{i,daily} \cdot (1 - \delta) + w_{i,two\ weeks} \cdot (\delta)$$

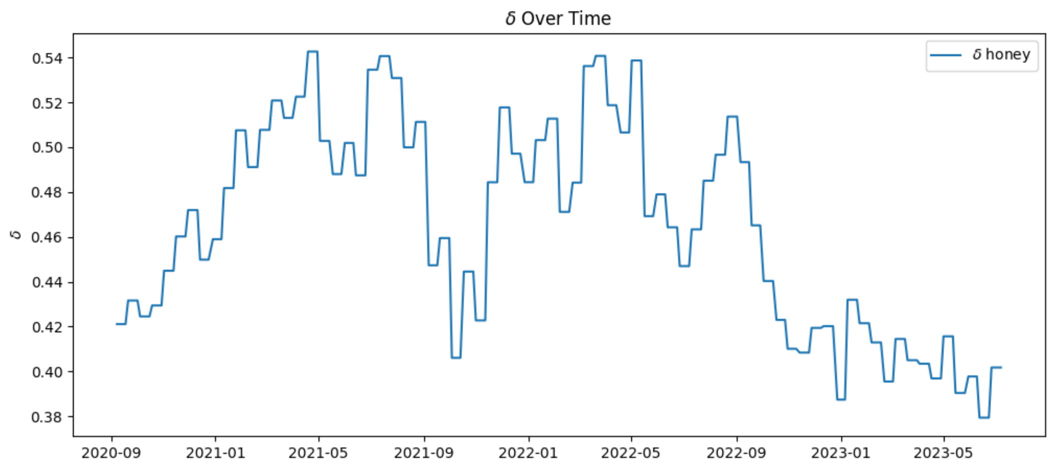
$$\sigma_p = \sqrt{w^T \cdot M_{shrunk} \cdot w}$$

$$\delta = \frac{\sigma_{daily} \cdot \sqrt{\vartheta}}{\sigma_{two\ weeks} \cdot \left(\sigma_{daily} \cdot \sqrt{\vartheta}\right)}$$

Where ϑ represents the number of trading days until the next rebalance. Converting δ as the proportional weight of volatility where the larger weight is on the lower volatility asset reduces the serial autocorrelation error of the daily return samples, while holding onto the momentum trends from the short-term timeframe weights compared to the longer-term weights. Additionally, since the period of prediction forward to convert daily volatility to a two-week volatility is small (≤ 10 trading days) Diebold, Hickman, Inoue, and Schuermann (1998) find it acceptable to scale risk by the square root rule to scale daily volatility.

The two portfolios’ volatilities are fairly balanced when comparing over time, however there are noticeable shifts as shown in Figure 2. Surprisingly over the time horizon of three years, the maximum value for δ is 0.543, while the minimum over the same time horizon is 0.379.

Figure 2 – Delta (Two Week Honey vs Daily Honey) Over Time



Results

Since the weight prediction using a covariance matrix is rolling and forward looking, it does not require a split between training and testing sets as was needed with the commodity optimization. However, since the underlying assets that are used in the portfolio are the 'optimized' commodity roll times, only the testing period of the commodities were used. The 'Mkt' and 'LSC' results can be found in Appendix C – Mkt-LSC Returns. The sample covariance matrix weight estimation results are in Figure 3, followed by the 'Honey' results in Figure 4, and finished with the combination results for all methods in Figure 5.

Figure 3 – Sample Covariance Weights

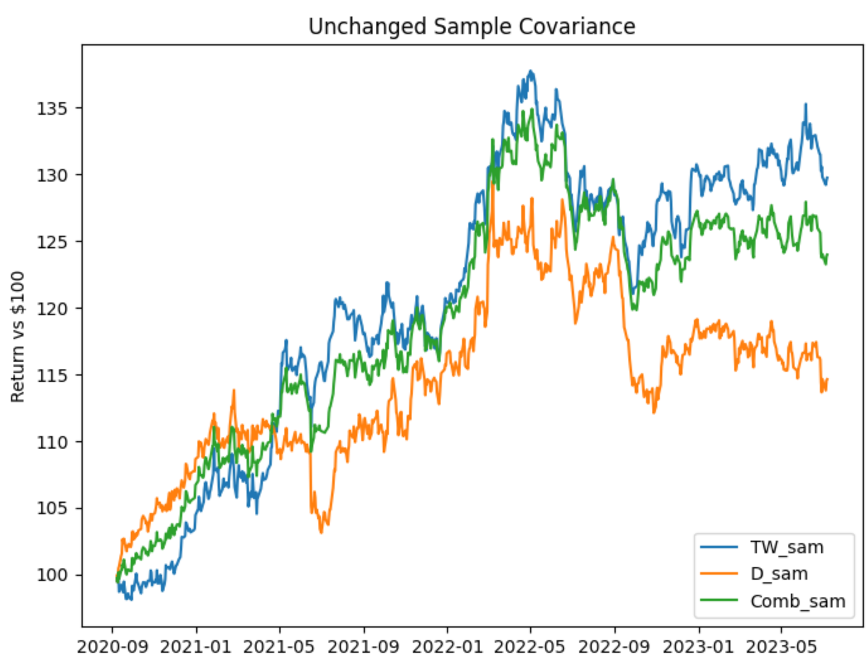


Figure 4 – Honey Covariance Weights

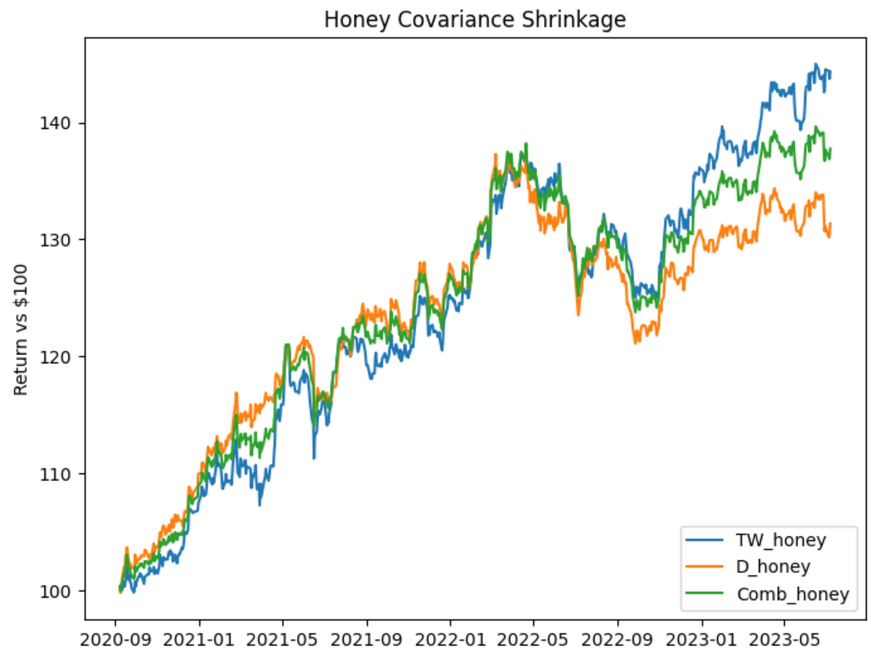
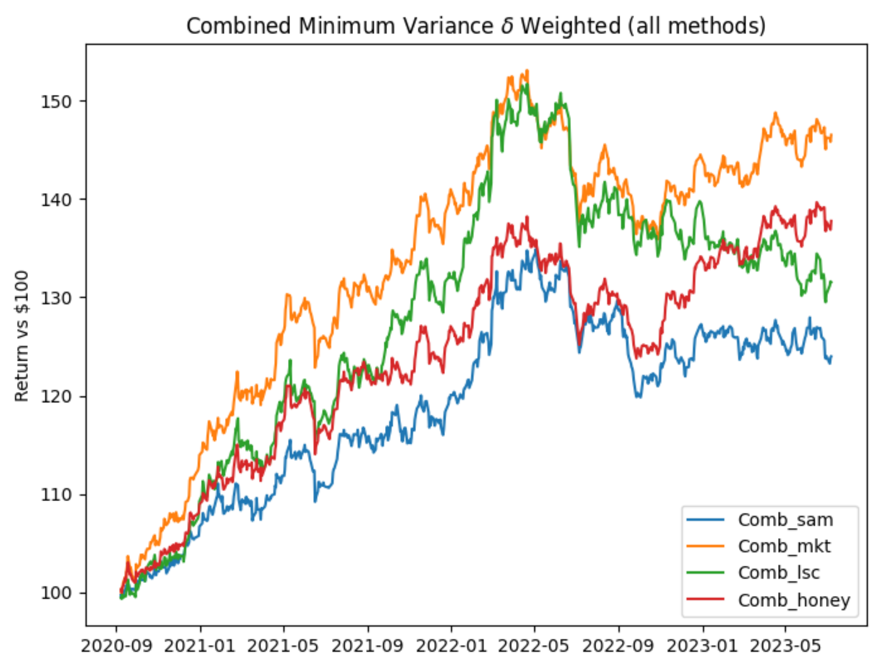


Figure 4 – Combined Weights, All Methods



Daily Return Standard Deviation * 100				
	Honey	Sample	Market	Large-Scale Covar
Combination	0.519	0.542	0.546	0.619
Two Weeks	0.569	0.674	0.615	0.674
Daily	0.553	0.636	0.569	0.613

The combined strategy which uses the minimized proportional weighted volatility measured by δ outperforms the individual timeframe standard deviations in all cases except the LSC model, where it marginally trailed the daily standard deviation of returns. Additionally, when compared on a standard deviation basis, the Honey model outperforms the other three methods tested. This fits with the rational that in a single asset class diversified portfolio, the Honey model which shrinks eigenvalues towards the mean of constant correlation. As previously mentioned, the goal of the GMV portfolio is not to predict the expected returns for assets, but to pick the portfolio that delivers the minimized variance in returns. In this case that is the Honey covariance shrinkage strategy with the combined weights.

Conclusion

Despite having the least optimal continuous timeline of prices from the testing period of the commodity rolling exploration, utilizing modern techniques of covariance matrix estimation, we can reduce the risk of the portfolio significantly. The process of covariance matrix shrinkage exploits common properties within financial return data to reduce estimation error between the sample and population covariance matrix. Additionally, with the novel development of a weighted minimum volatility measurement for weights, we are able to reduce the risk of the portfolio further by setting a proportional weight of weights on the lower of the two-time frames (two-week vs daily) returns.

In the future, I would like to bring the covariance matrix shrinkage analysis to new asset classes, as well as potentially look at the impact on weights optimizing individual asset classes will have on a multi-asset class portfolio if each individual asset class is calculated to a GMV portfolio. Additionally, I would like to explore dynamically rolling contracts for commodities forward since the results of the empirical analysis found the vast underperformance of a static roll date/days strategy. Lastly, I did not get a chance to explore non-linear covariance matrix shrinkage which has recently been pushing the frontier of portfolio optimization research, and it would be interesting to explore further.

References

Diebold, F.X. Hickman, A., Inoue, A. and Schuermann, T. (1998). Converting 1-Day Volatility to h-Day Volatility: Scaling by Root-h is Worse than You Think. *Warton Financial Institutions Center,* Working Paper 97-34

Large, S., Reeves, B., and Haji, A. (n.d.). Revisiting Shrinkage: Covariance Matrix Estimation in the Modern Age of Portfolio management. *Viewpoint Investment Partners.* Retrieved November 3, 2023 from https://covariance.viewpointexperience.ca/

Ledoit, O., and Wolf, M., (2003). Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *Journal of Empirical Finance,* Vol.10, 603-620

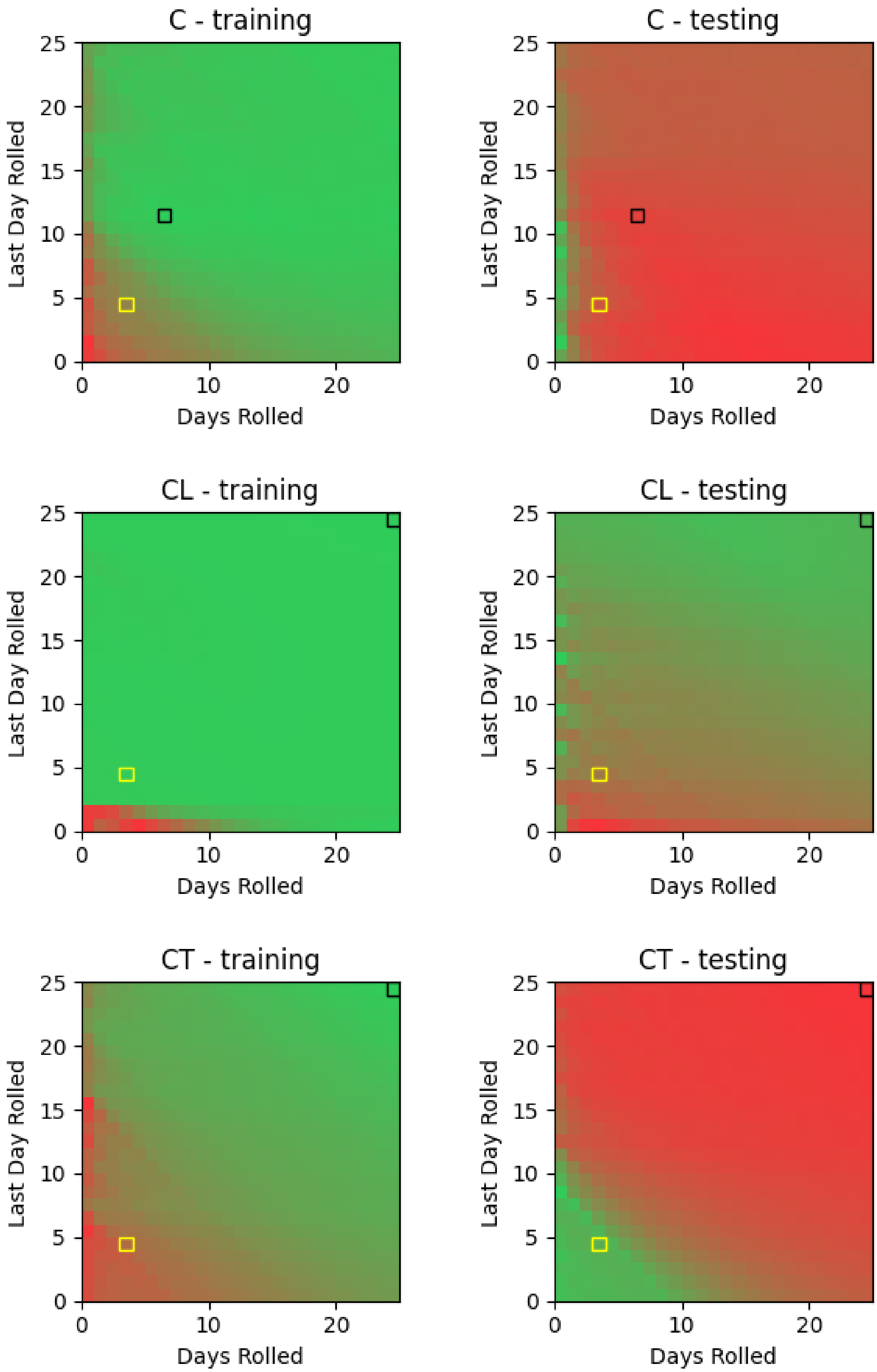
Ledoit, O., and Wolf, M., (2004a). Honey, I Shrunk the Sample Covariance Matrix. *Journal of Portfolio Management.* Vol.30, 110-119

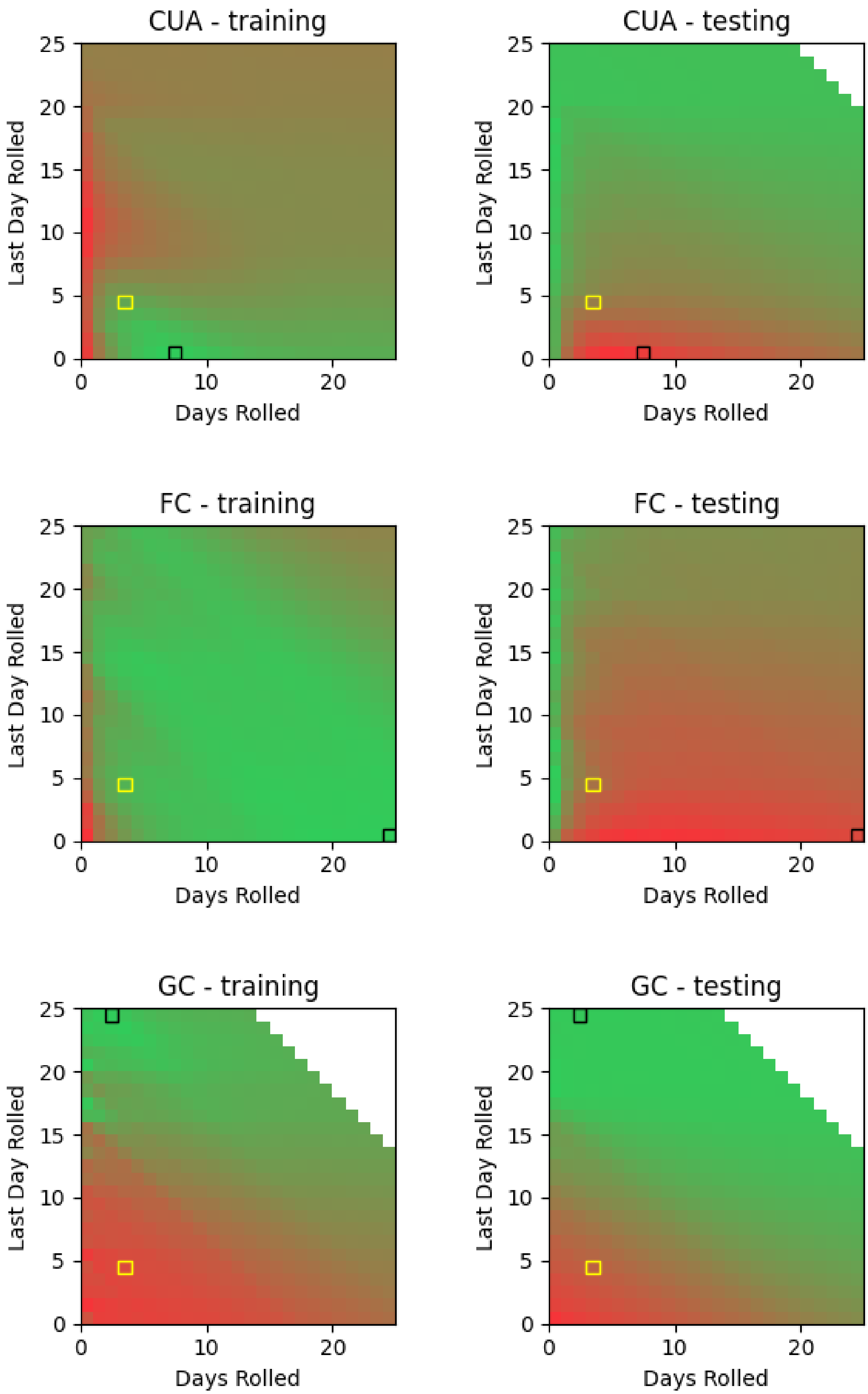
Ledoit, O., and Wolf, M., (2004b). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis.* Vol.88, 365-411

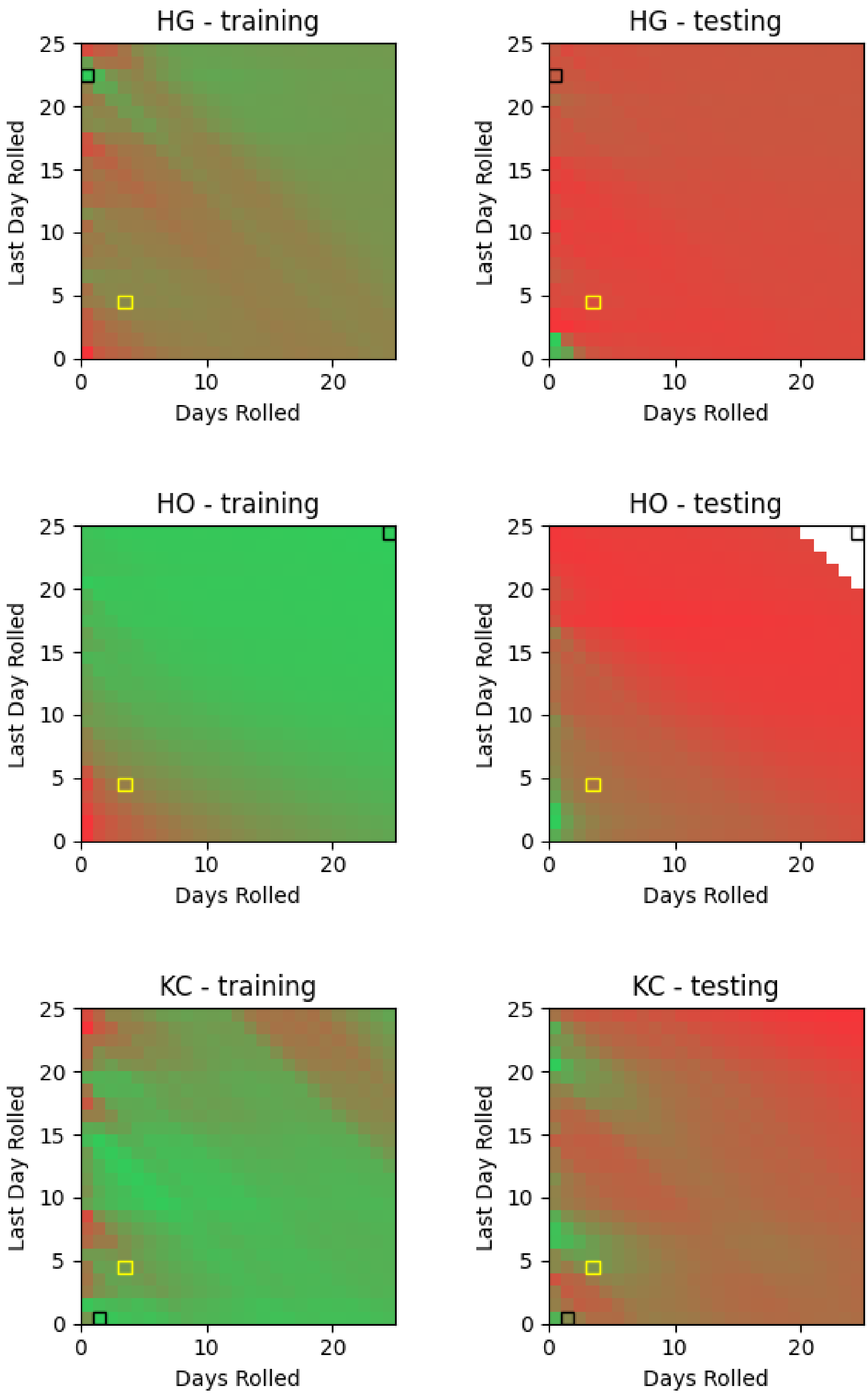
Ledoit, O., and Wolf, M., (2020). The Power of (Non-)Linear Shrinking: A Review and Guide to Covariance Matrix Estimation. *Journal of Financial Econometrics,* Vol.20, No. 1, 187-218

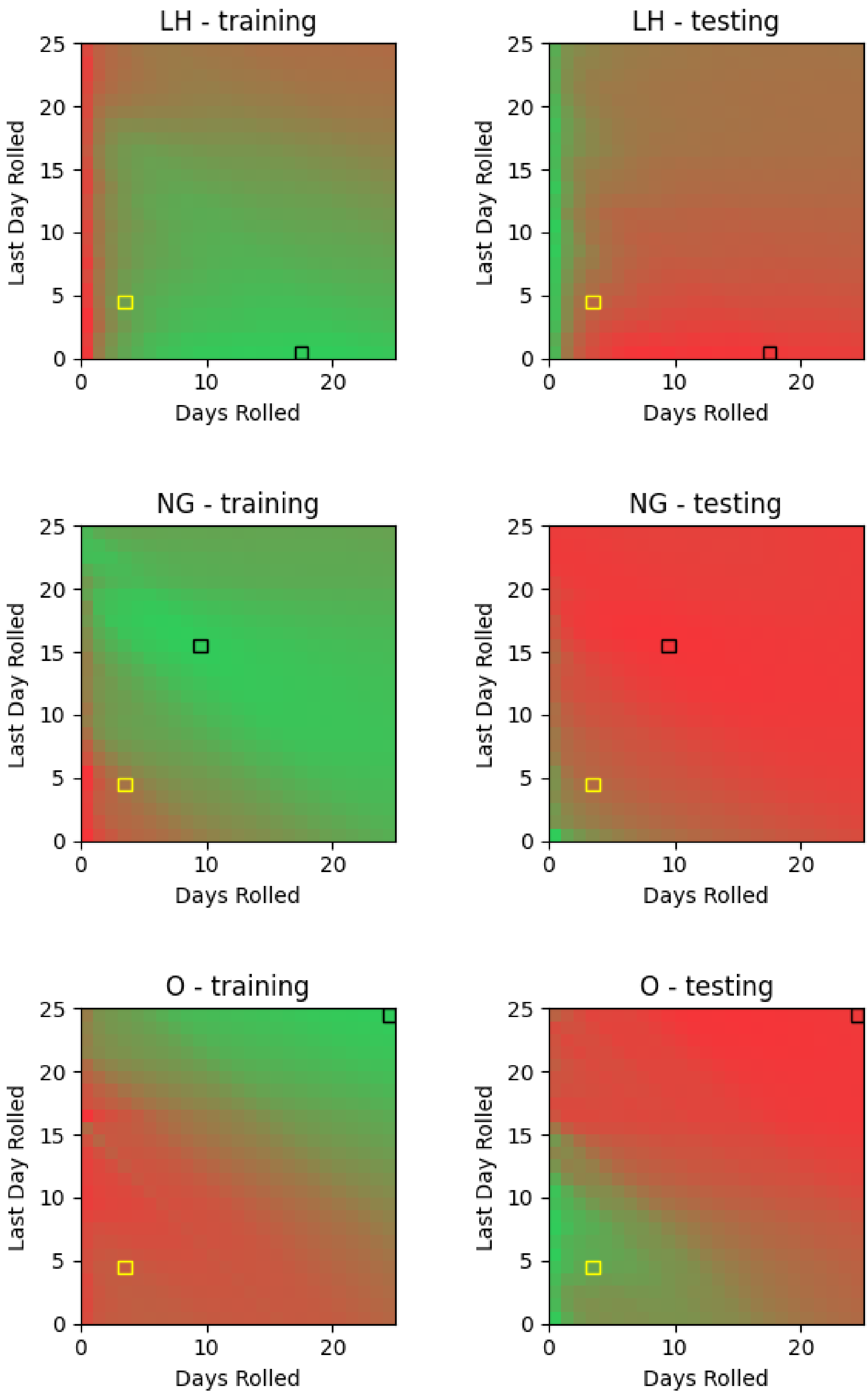
Taylor, N., (2016). Roll Strategy Efficiency in Commodity Futures Markets. *Journal of Commodity Markets*, Vol. 1 No. 1, 14-34

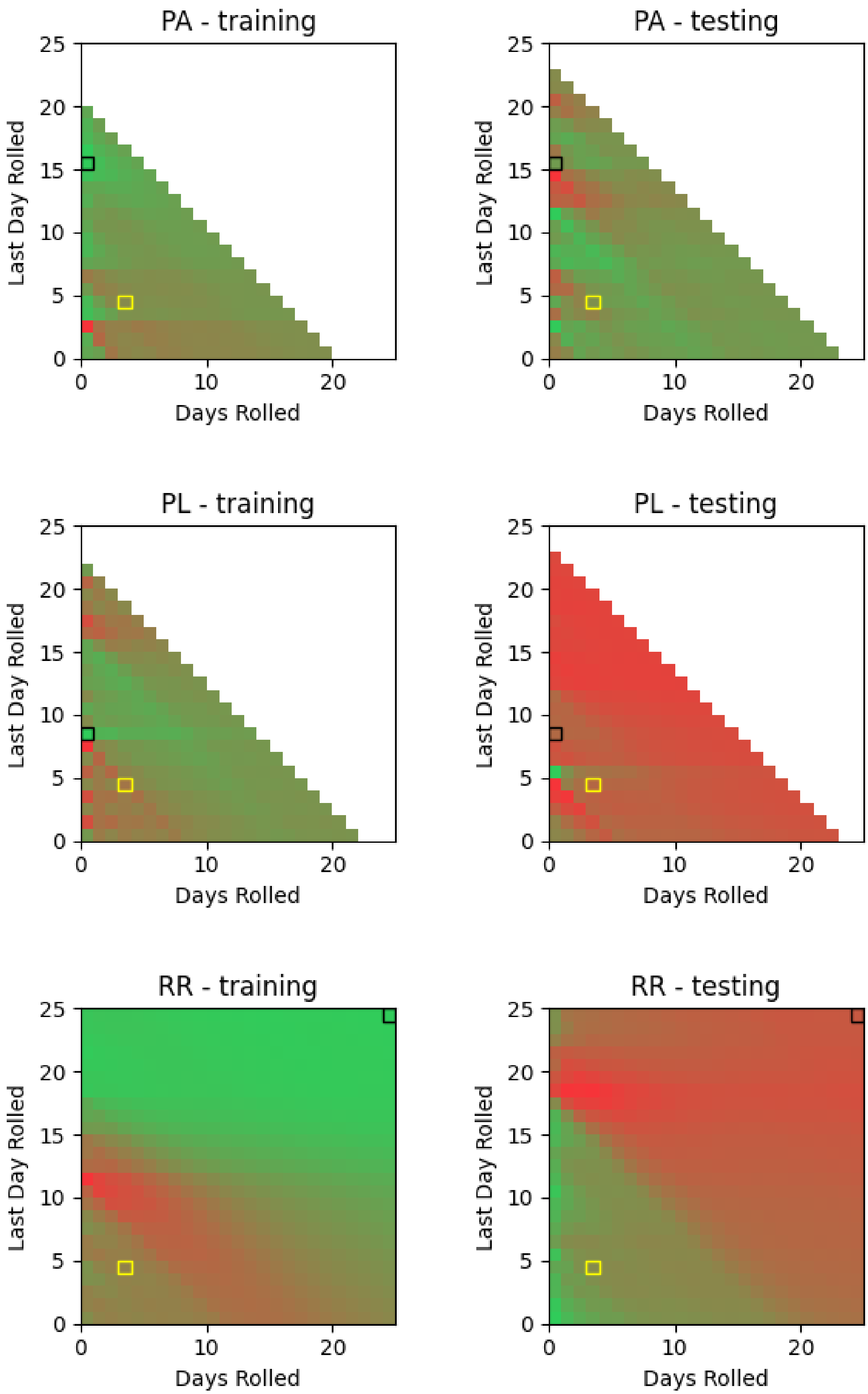
Appendix A – Commodity MVO Plots

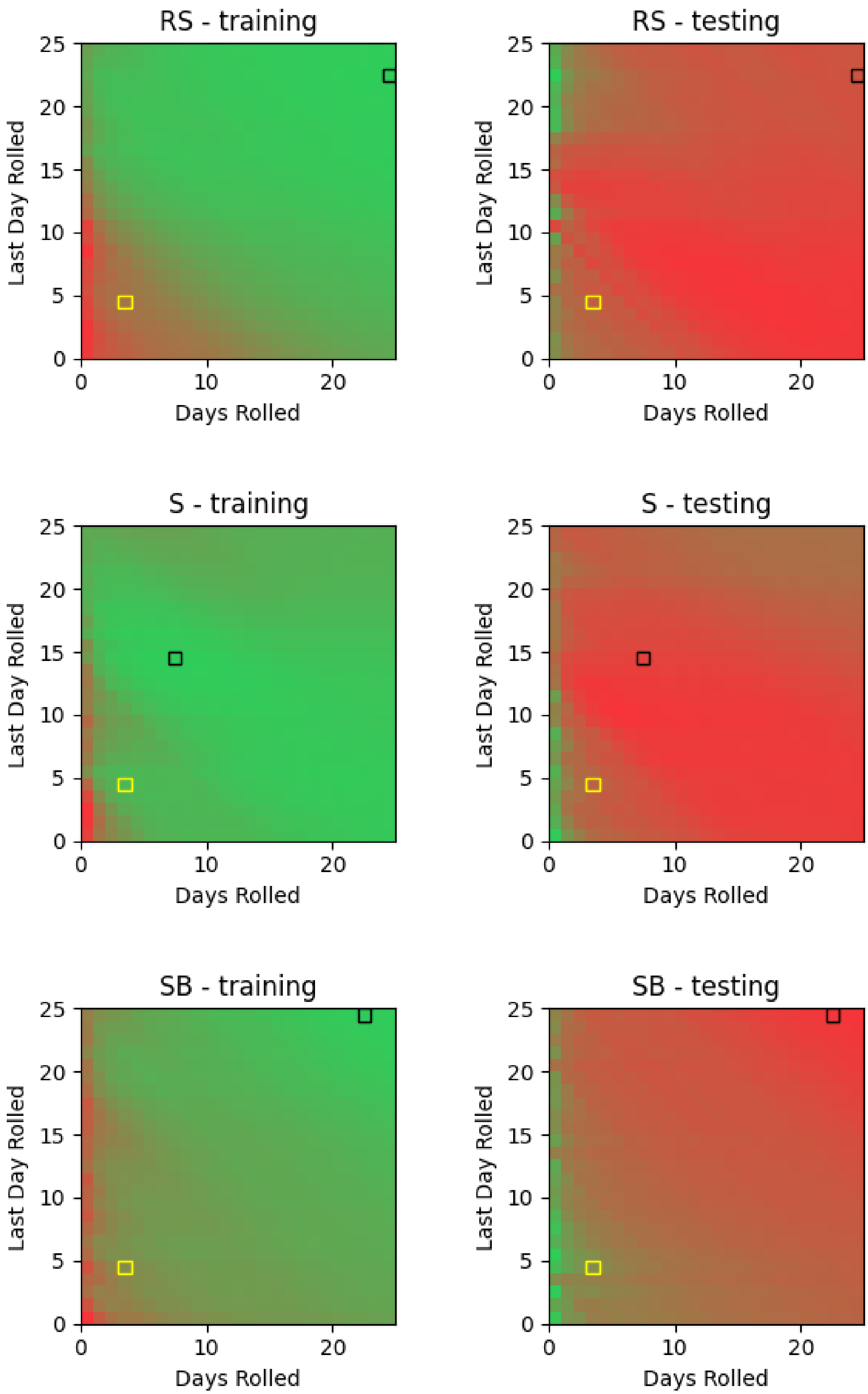


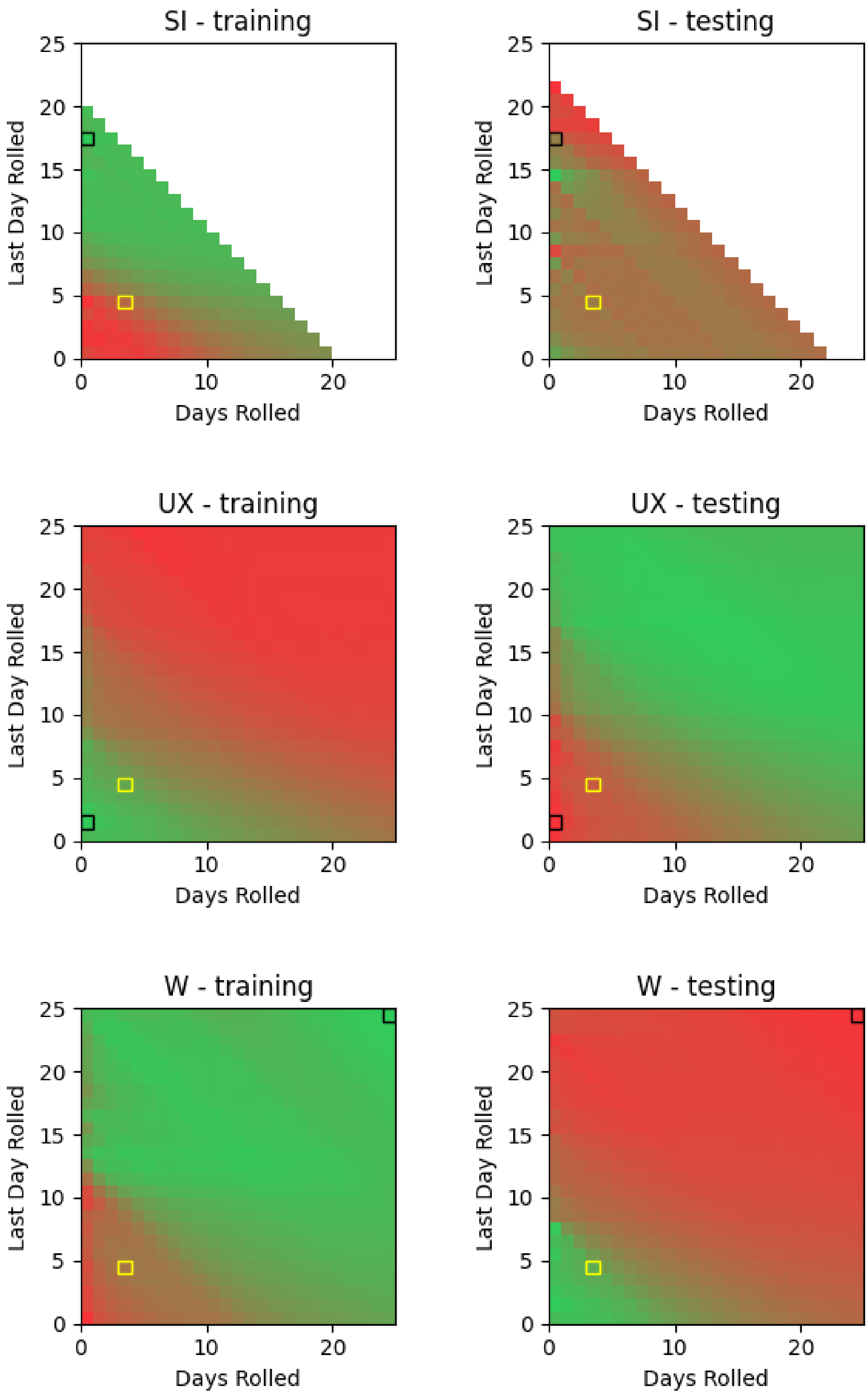

CUA - training
CUA - testing
FC - training
FC - testing
GC - training
GC - testing
Last Day Rolled
Days Rolled
0
5
10
15
20
25


HG - training
HG - testing
HO - training
HO - testing
KC - training
KC - testing
Last Day Rolled
Days Rolled
0
5
10
15
20
25


LH - training
LH - testing
NG - training
NG - testing
O - training
O - testing
Last Day Rolled
Days Rolled
0
5
10
15
20
25
0
10
20


PA - training
PA - testing
PL - training
PL - testing
RR - training
RR - testing
Last Day Rolled
Days Rolled
0
5
10
15
20
25


RS - training
RS - testing
S - training
S - testing
SB - training
SB - testing
Last Day Rolled
Days Rolled
0
10
20
5
15
25

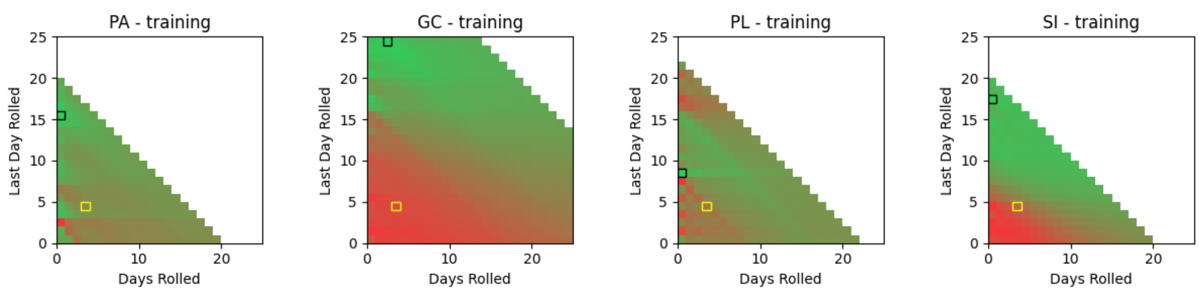

SI - training
SI - testing
UX - training
UX - testing
W - training
W - testing
Last Day Rolled
Days Rolled
0
5
10
15
20
25

Appendix B – Contract Liquidity Constraints

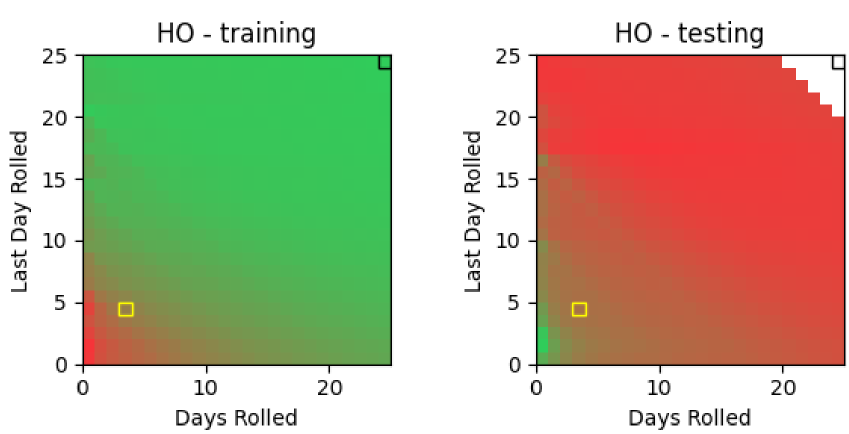
The optimization sample size for the metals Silver, Platinum, Palladium, and Gold were significantly shorter than the rest of the commodities. This is due to a combination of the three following issues with the contract data.

1. There is a lack of liquidity in the contracts meaning that these contracts suffer from the same problems as thinly traded securities as explored in assignment 1.
2. The contract matching code pads the last traded day price if Bloomberg is missing a datapoint. (which you cannot trade since there is likely a volume of 0 on that day)
3. There is not a long enough overlapping period for the data selected.

The first two Silver contracts have 22 days of overlap between the current month and the next month's contract expiry. This means that to have an accurate sample for a continuous roll, the number of roll days plus the last roll date must be less than or equal to 22 days. However, setting a limit on the number of overlapping contract days changes month to month with liquidity. Additionally, the first day that the November 2013 contract traded for Silver was September 30th and had a trading volume of 1. A single contract traded is not considered an efficient market. Platinum, Gold, and Palladium all suffer from similar issues when investigated further.

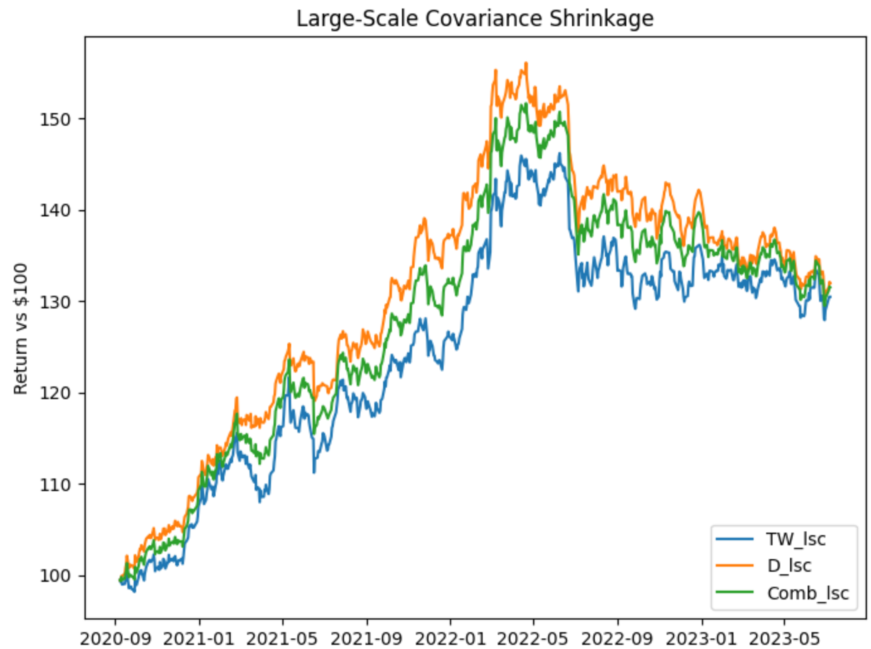


Further, a similar issue happened during the testing period for Heating Oil, where it's optimized mean variance location was out of historically tradable sample ranges. The materiality of this issue is low since if a trader were trading the MVO strategy live, they would simply wait until the first day and spread out the number of contracts that they need to roll over a shorter number of days to keep the last DTE consistent.



Appendix C – Mkt-LSC Returns

Largs-Scale Covariance Matrix Shrinkage Performance



Market Model Covariance Matrix Shrinkage Performance