EMPIRICAL ANALYSIS OF OPTIMIZATION ALGORITHMS FOR PORTFOLIO ALLOCATION

By

Andrew Bolin

A Thesis Submitted to the Graduate

Faculty of Rensselaer Polytechnic Institute

in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

Major Subject: COMPUTER SCIENCE

Examining Committee:

Malik Magdon-Ismail, Thesis Adviser

Boleslaw Syzmanski, Member

Aparna Gupta, Member

Rensselaer Polytechnic Institute
Troy, New York

April 2013
(For Graduation May 2013)

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGMENT

I am incredibily grateful to my adviser Malik Magdon-Ismail. When circumstances forced me to find a new project on short notice, professor Magdon-Ismail provided me not only with a new project, but also one that perfectly aligned with my interests. It has been a pleasure working with him and the rest of Team Malik.

ABSTRACT

Portfolio optimization algorithms were tested using historical S&P100 data. A traditional Mean-Var algorithm is tested as well as two alternative risk methods. The alternative risk methods used the maximum drawdown (MDD) as the measure of risk rather than the standard deviation of returns. The long term performance of each portfolio produced by these algorithms was compared to various benchmarks. It was found that the algorithms often outperformed the benchmarks in real rate of return. However, when return was adjusted for risk, the algorithms generally underperformed the benchmarks. A hierarchical method involving clustering stocks before allocation was also tested, but was found to underperform allocation without clustering. MDD-minimization subject to a return constraint was found to be the only optimization algorithm that outperformed the benchmarks in each measure of performance. Closer examination revealed that this algorithm closely matched the benchmarks during periods of continuous market growth, but significantly outperformed them during periods of continuous market decline. This result supports the use of maximum drawdown as an alternative measure of risk in portfolio optimization when investors are assumed to be risk-averse.

1. INTRODUCTION AND HISTORICAL REVIEW

Although formal mathematical models for diversification are relatively new, diversification has been practiced by investors, merchants, farmers, and more throughout history as a way to reduce the risks they face. Harry Markowitz, who is widely considered to be the father of modern portfolio theory, stated that diversification of investments long predated his work on portfolio selection (Markowitz, 1999). He even quotes Shakespeare to prove his point:

My ventures are not in one bottom trusted,
Nor to one place; nor is my whole estate
Upon the fortune of this present year:
Therefore my merchandise makes me not sad.
(*Merchant of Venice*, Act 1, Scene 1)

Written at the end of the sixteenth century, this quote clearly presents an understanding of diversifcation long before Markowitz's seminal work. Even earlier examples of diversifcation also exist. Ancient Chinese merchants would divide their shipments among a fleet of junks. If one ship sank, only a fraction of the shipment would be lost. In the fourth century, Rabbi Isaac Bar Aha proposed that "One should always divide his wealth into three parts: a third in land, a third in merchandise, and a third ready to hand" (Tractate Baba Mezi'a, folio 42a). Even though these heuristic and naive methods of diversification have existed for centuries, only in the last sixty years have formal models for diversification been extensively developed and studied.

1.1 Modern Portfolio Theory

The overall goal of modern portfolio theory is to maximize risk adjusted return. Adjusting for risk is desirable as investors have long been shown to be risk averse (Menezes & Hanson, 1970; Pratt, 1964; Tobin, 1958). Using mathematical models of risk and diversification, investors attempt to optimize their portfolios to produce the greatest risk adjusted return. Various representations and estimations of risk exist resulting in various methods of optimization. In this research, the traditional Markowitz Mean-Variance, and two MDD optimization methods are examined.

1.1.1 Markowitz or Mean-Variance Model

Harry Markowitz was not the first to show that diversification can reduce risk, but he was among the first to develop a formal model for reducing risk without reducing return. He showed that portfolios constructed from risky instruments can be mapped in a space of expected return versus risk. When all possible portfolios fall in this space, those portfolios which minimize risk at given levels of expected return, or maximize return at a given level of risk, can be identified. These portfolios are called efficient, because no other portfolio will offer the same return with less risk or more return with the same risk. This model is considered to be a two-moment decision model as the allocation is decided based on the first and second moments of the portfolio's returns. Not only did Markowitz establish this model, he also showed how to construct a quadratic program capable of finding efficient portfolios when the second central moment, the variance, of the portfolio's returns is used as the measure of risk (Markowitz, 1952). This quadratic program is discussed in the following chapter and used in this paper's experiments.

1.1.2 MDD Models

The maximum drawdown of an investment (MDD) is considered to be an alternative measure of risk in modern portfolio theory. MDD is defined as the maximum peak to trough difference in the cumulative return series of an investment over a specific time horizon. This measure is especially important to fund managers as large draw downs are unacceptable to investors. This makes maintaining a small MDD critical to the survival of most managed funds. The use of MDD as a risk measure has become more attractive as recent studies have suggested that returns are positively skewed (Defusco et al., 1996; Simkowitz & Beedles, 1978; Singleton & Wingender, 1986). The Mean-Variance method assumes that returns are normally distributed. Positively skewed returns disrupt this assumption leading to suboptimal portfolio allocations. Many researchers are taking this issue seriously and the current trend is to develop new higher moment optimizations (Jondeau & Rockinger, 2006). Rather than develop optimizations of higher moments, it is possible that researchers should simply optimize using a different risk measure. The MDD of a portfolio depends on the mean and volatility of its returns; however, simulation experiments have shown that the MDD depends very little on skewness or kurtosis (Burghardt et al., 2003). A possible explanation for the observations made in these experiments is provided by the author. Drawdowns are constructed by summing over a sequence of returns. Considering the central limit theorem, adding returns drawn from a skewed distribution is expected to produce a sample that is closer to normally distributed (Burghardt et al., 2003). If these results hold true, then an optimization using MDD as the measure of risk should produce a superior allocation than the flawed Mean-Variance optimization. An analysis by Hamelink and Hoesli shows that portfolios optimized using MDD, rather than standard deviation, as the measure of risk result in a lower MDD and only a slightly larger standard deviation for the same level of return (Hamelink & Hoesli, 2004). An additional benefit of using MDD as the risk measure is that the optimization problem can be solved using a linear program instead of a quadratic program. This makes it much simpler to solve both sides of the allocation problem: minimizing risk subject to a return constraint, and maximizing return subject to a risk constraint. Minimizing MDD subject to a return constraint and maximizing return subject to an MDD constraint will be the remaining methods tested in this research and will be discussed in the next chapter.

1.1.3 Contributions

This research provides new empirical data for the Mean-Var method by calculating the Calmar ratios produced by the method. Some of the first empirical results for the MDD Constraint and Min-MDD methods are presented in this research. Additionally, optimal choices for risk constraints and length of history for each method are determined and presented in Table 5.1.

2. METHODS

2.1 Selection Algorithms

The general problem is, given N instruments and a history of their prices produce a portfolio using those instruments which maximizes risk-adjusted return. For each method, the output is Θ, a $N \times 1$ vector of portfolio weights with $\sum_{i=1}^{N} \Theta_i = 1$. In general, it is possible for weights to be negative which represents the short selling of a stock. However, this research ignores short selling and enforces the the condition that $\forall_i \Theta_i \geq 0$. Each method, besides the uniform method, is given a $T \times N$ matrix of prices as input. This matrix contains the price for each of the N instruments on T different days. These T days will be sequentially ordered, and the $T \times N$ matrix of prices will represent a window of pricing history. The different selection algorithms build their portfolios using the prices in the $[0, T]$ time horizon. The expected return of the portfolio is $\mu^T \cdot \Theta$, where μ^T is the vector containing the average return for each usable stock. The MDD methods can be found in standard form in Appendix A.

2.1.1 Uniform

This is the simplest method to create a diversified portfolio and will be used as one of the benchmarks to compare the other methods against.

$$\text{for all } i = 1, .., N, \quad \Theta_i = \frac{1}{N} \tag{2.1}$$

Using this method, an equal amount of capital is allocated to each stock during each trading period. Allocating capital in this manner protects the portfolio from sudden crashes in individual instruments, but also hinders large gains from overachieving instruments.

2.1.2 Volatility Weighted

This method is similar to the uniform method as capital will be allocated to each usable stock in the universe. However, the amount of capital allocated to each stock will be proportional to the inverse of the volatility of the stock's returns in the recent past.

$$\forall_i \Theta_i = \frac{\frac{1}{\sigma_i}}{\sum_{i=1}^{N} \frac{1}{\sigma_i}} \tag{2.2}$$
$$\sigma_i = std\_dev(R_i)$$

Volatility is measured using the standard deviation of the stock's returns and is reprsented by the symbol σ. Using the $T \times N$ matrix of prices, the $T \times N$ matrix of returns for each stock over the $[0, T]$ time horizon is constructed. This matrix is denoted R, and R_i is the column vector of returns for stock i. Each volatility quotient is divided by the sum of all volatility quotients so as to maintain the $\sum_{i=1}^{N} \Theta_i = 1$ condition. This method serves as the second benchmark against which the other methods will be compared.

2.1.3 Mean-Variance Method

Σ is the covariance matrix of the returns for the usable instruments in the universe. $\Theta^T \Sigma \Theta$ is the variance of the returns of the portfolio described by Θ. This method minimizes the risk factor subject to a return constraint. The return constraint used in this research is a weighted average of the minimum and maximum average returns amongst all usable instruments.

$$\begin{aligned} \min_{\Theta} \Theta^T \Sigma \Theta, \\ s.t. \\ \mu^T \cdot \Theta \geq \lambda \cdot max(\mu) + (1 - \lambda) \cdot min(\mu), \\ \forall_i \Theta_i \geq 0 \\ \sum_{i=1}^{N} \Theta_i = 1 \end{aligned} \tag{2.3}$$

2.1.4 Maximizing Return Subject to an MDD Constraint

Chekhlov et al. (2005) provides an example of maximising expected return subject to various draw down constraints and provides several empirical results using a universe of 32 instruments. The major distinctions between this previous work and the work presented in this paper is the method for selecting the MDD constraint, the number of instruments in the universe, and the types of instruments in the universe. In Chekhlov et al. (2005), the constraint is defined as a fraction of the capital value of the portfolio, and the instruments used are from a variety of liquid markets including currency markets, equity indices, and precious metals. In this research the constraint is simply the MDD of a uniform portfolio constructed over the same $[0, T]$ time horizon, and the universe is composed of all stocks that were members of the S&P100 at the time of the portfolio's construction.

$$\begin{aligned} \max_{\Theta} R_T \cdot \Theta, \\ s.t. \\ MDD(R_{\Theta}) \leq \lambda \cdot MDD(Uni), \\ \forall_i \Theta_i \geq 0 \\ \sum_{i=1}^{N} \Theta_i = 1 \end{aligned} \tag{2.4}$$

2.1.5 Minimizing MDD Subject to a Return Constraint (Min-MDD)

The MDD Constraint method maximizes return for a given level of risk while this method minimizes the risk factor for a given level of return. This makes it more comparable to the Mean-Var method as they both attempt to construct efficient portfolios by minimizing risk. Similar to the Mean-Var method, this method also uses a weighted average of the minimum and maximum average returns to determine the target level of return.

$$\begin{aligned}\min_{\Theta} MDD(R_{\Theta}),\\ s.t.\\ \mu^T \cdot \Theta \geq \lambda \cdot max(\mu) + (1-\lambda) \cdot min(\mu),\\ \forall_i \Theta_i \geq 0\\ \sum_{i=1}^{N} \Theta_i = 1\end{aligned} \tag{2.5}$$

A benefit of this method is that comparing it to the Mean-Var method is simple as both methods can be tested using the same return constraint.

2.2 Clustering

Clustering is used to group highly correlated stocks and to ensure that the correlation between separate groups is as low as possible. Diversification is accomplished by allocating within groups, followed by allocating between groups. As separate groups are constructed to be uncorrelated, diversifying between groups should reduce risk.

2.2.1 Clustering Algorithm

First, the $1 \times T$ performance vector A is contrsucted for each stock. For stock i, A_i is built using the $1 \times T$ vector of daily returns for stock i, $R_i^{'}$.

$$\text{for } j = 1..T, \quad A_{ij} = \begin{cases} 1 & \text{if } R_{ij}^{'} > \mu(R_{ij}^{'}) + \sigma_{R_i^{'}} \\ -1 & \text{if } R_{ij}^{'} < \mu(R_{ij}^{'}) - \sigma_{R_i^{'}} \\ 0 & \text{o.w.} \end{cases} \tag{2.6}$$

$$\begin{aligned}&\text{where,}\\ \sigma_{R_i^{'}} &= std\_dev(R_i^{'}),\\ \mu(R_i^{'}) &= mean(R_i^{'})\end{aligned}$$

Next, a stock s_i is selected at random to serve as the first center, c_1. The correlation between its returns and the returns of each other stock is computed, using their corresponding A vectors. This $1 \times N$ vector of correlations will be the first row in a $k \times N$ matrix of correlations, ρ, where k = number of clusters. Until k centers have been chosen, the stock most uncorrelated with the previous center becomes the next center. The next row of the correlations matrix is built once the new center is determined. Once the correlations matrix is filled, ρ_i will be the vector of correlations between stock i and each center. For each stock i, $i = 1..N$, i is placed in the cluster with center c_j, where j satisfies $\rho_{ij} = max(\rho_i)$.

3. DATA AND EXPERIMENTS

3.1 Data

The closing stock prices for 1220 publicly traded companies from 12/30/1983 to 12/15/2011 were compiled from historical stock market data. Along with the daily closing prices for each stock, binary indicators for membership in the S&P 500 and the S&P 100 were also compiled. Membership in the S&P 500 was tracked for the entire time period, and membership in the S&P 100 was tracked from 09/11/1989 to the end of the period. The period from 09/11/1989 to 12/15/2011 spans 5614 trading days, giving a history of 1220×5614 prices. The 1220×5614 matrix of prices and equivalently sized matrix of S&P 100 indicators serve as the input for the simulations.

3.2 Simulations

Each method is tested by simulating investing in the portfolios constructed by the method using the S&P 100 or S&P 500 data. Simulations begin with a capital investment of 1, and shares are assumed to be infinitely divisible. To simulate long-term investing, portfolios are constructed, held, and rebalanced multiple times throughout the simulation. Each constructed portfolio is held for a predetermined number of trading days and then rebalancing occurs. To rebalance, the chosen allocation algorithm is used to create a new set of portfolio weights using the current trading day as the starting time. Figure 3.1 shows how rebalancing creates a moving history window. *nhor* is the length of the period in which the portfolio is held, and *phor* is the length of the history given to the allocation algorithm. When $phor > nhor$, there is a gradual change in the history provided to the allocation algorithms. When the first allocation occurs, the history is composed of the last *phor* days. At the next allocation, the history loses the oldest *nhor* days and gains the newest *nhor* days. This change occurs at every allocation except for the first. Rebalancing is critical to ensure that the portfolio remains in the efficient set despite changing market conditions. When portfolios are rebalanced, the captial available

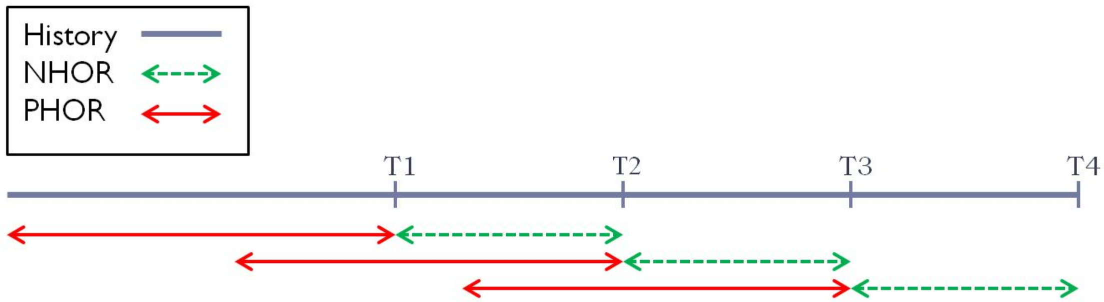


Figure 3.1: Rebalancing and the Moving History Window

is decreased by 3 basis points to simulate transaction costs.

The capital worth of each constructed portfolio is recorded for each day in the simulation. Records for each method are used to calculate the average daily return, average annual rate of return, and average annual MDD when investing using that method. The risk-adjusted performance metrics discussed in the following chapter require these calculations.

3.3 Simulation with Clustering

Simulations with clustering require a few additional steps. At the beginning of each trading period, the usable stocks are clustered into groups using the clustering algorithm. For each cluster, a vector of weights is produced using one of the selection algorithms. This vector of weights will be γ. The length of γ will be equal to the size of the cluster it was produced from, and $\sum_i \gamma_i = 1$. After portfolios have been constructed from each cluster, meta-stocks representing each cluster are built. A meta-stock is essentially a representative of a cluster. It takes the $T \times M$, where M is the number of stocks in the cluster, history of prices for the cluster, and consolidates them into a single $T \times 1$ history for a single stock. Let C be the $T \times M$ history of prices for the cluster and S be the $T \times 1$ meta-stock history. The price history of the meta-stock is weighted using γ.

$$\text{for i=1..T} \quad S_i = \sum_{j=1}^{M} \gamma_j \cdot C_{ij} \tag{3.1}$$

Using the price histories for each meta-stock as input, an allocation algorithm is used once again to produce a set of weights denoted ω. The final weight for each stock in Θ is found by multiplying the stock's corresponding γ by its cluster's weight in ω. The portfolio described by Θ is then held for *nhor* days before rebalancing, at which point the process is repeated.

4. RESULTS

Simulations were run for each algorithm with *nhor* set to 20 trading days and *phor* set to either 60, 120, 250, or 500 trading days. For the Markowitz and MDD-Minimization methods, $0.1, 0.2, \ldots, 0.9$ are used as values of λ. When maximizing return subject to an MDD constraint, $0.1, 0.2, \ldots, 2.0$ are used as the values of λ. Clustering was simulated for k=3 and k=5. All compiled results can be found in Appendix B.

4.1 Measuring Portfolio Performance

Three metrics for measuring portfolio performance are used to compare the different optimization methods, one raw measure, and two risk adjusted measures.

4.1.1 Rate of Return

The simplest method of measuring porfolio performance is to examine its real rate of return. This shows exactly how much return on investment a strategy provided during the simulation period. Due to the length of the simulation period, average annual rate of return is used. Several configurations of λ, *phor*, and optimization method produced returns well above the benchmarks. As seen in Figure 4.1, each optimization method produces greater returns than the benchmarks for several values of λ when $phor = 500$. However, a high rate of return does not guarantee that an investment was optimal. Figure 4.2 depicts the returns realized by the method with the highest rate of return and the method with the largest Sharpe ratio. While the method with the high Sharpe ratio produced a much lower return, it did so without wild gains or losses. The high return method experienced massive swings in gains and losses, losing more than fifty percent of its total value at several points. One method presents high risk and high reward while the other low risk and low reward. This example shows the desirability of adjusting returns for risk. Without adjustment, it is unclear which method should be chosen. After adjusting for the risk involved, the lower return method is to be preferred.

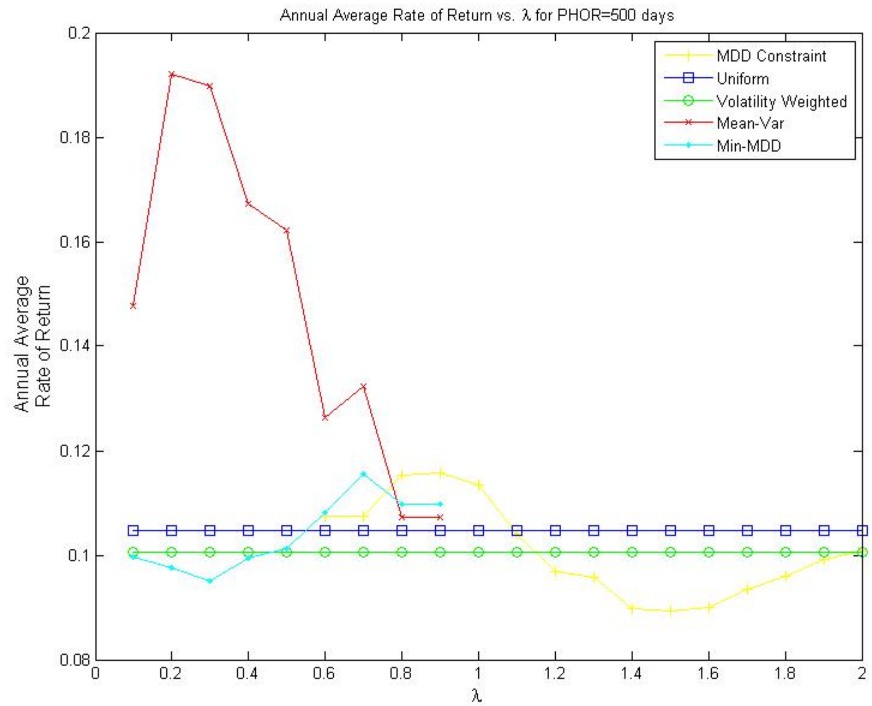


Figure 4.1: Average Annual Rate of Return when PHOR = 500 Days

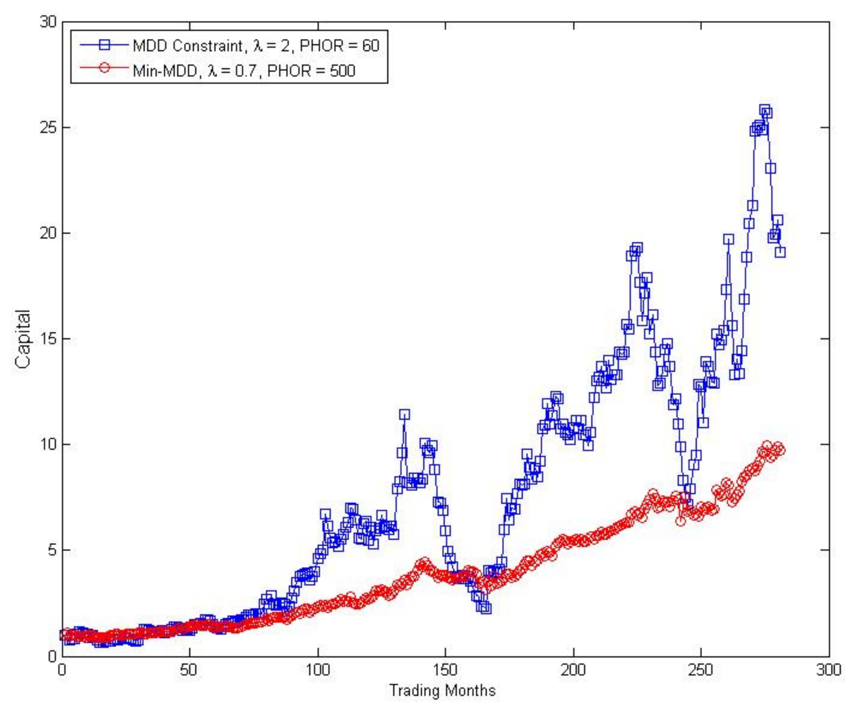


Figure 4.2: Highest Rate of Return vs. Largest Sharpe Ratio

4.1.2 Sharpe Ratio

Let R_D(T) = average daily return over [0, T], and σ(T) = standard deviation of daily returns over [0, T]. To compute an annualized Sharpe ratio, the ratio is multiplied by the square root of the number of trading days in a year.

$$Sharpe(T) = \frac{R_D(T)}{\sigma(T)} \cdot \sqrt{250} \tag{4.1}$$

The Sharpe ratio has long been used as a measure of risk adjusted performance as it rewards larger returns while penalizing large swings in those returns. Ideally, an investment will provide consistent large returns. Investments that approach this ideal will have high Sharpe ratios. While the optimization algorithms often outperformed the benchmarks in annual rate of return, they often generated lower Sharpe ratios than the benchmarks. This means that the optimization algorithms generally produced higher variance returns than the benchmarks, which is undesirable. Figures 4.3 and 4.4 clearly present this phenomenon. The annual rate of return for the MDD constraint method approaches twenty percent; however, its Sharpe ratio plummets as the rate of return increases. This suggests that the additonal return was gained by taking on a substantial amount of risk.

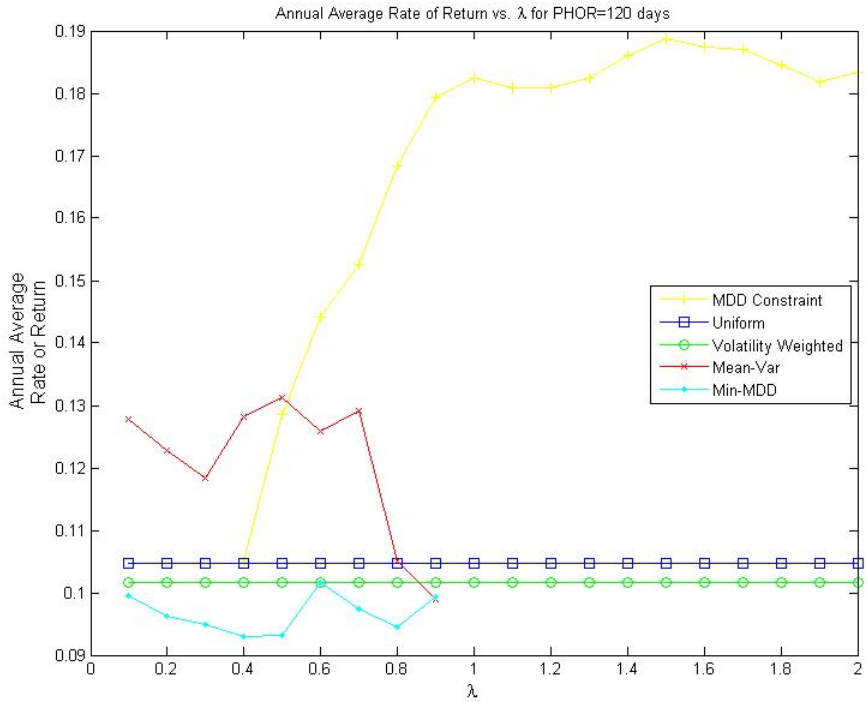


Figure 4.3: Average Annual Rate of Return when PHOR = 120 Days

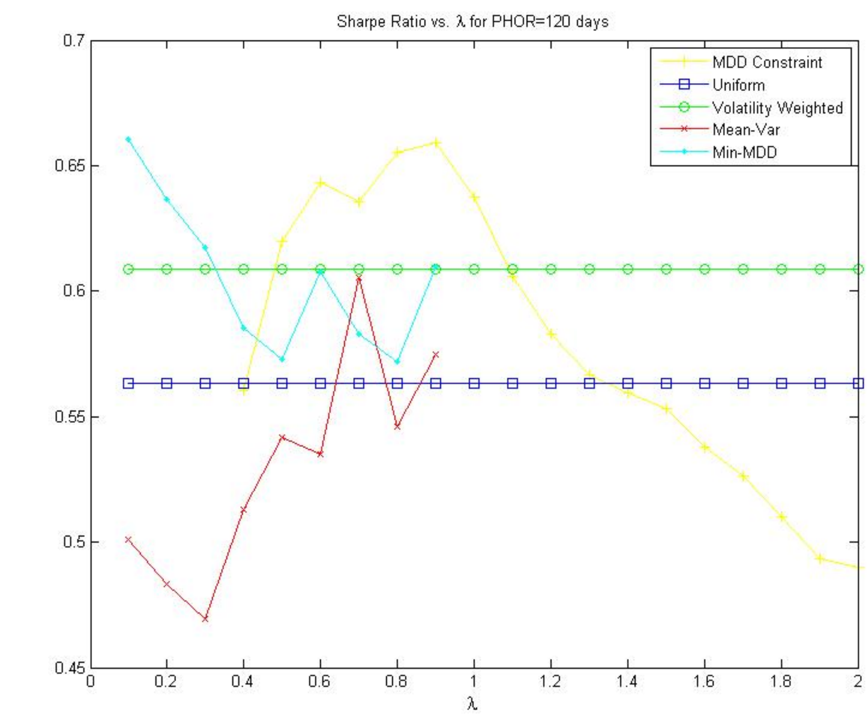


Figure 4.4: Sharpe Ratio when PHOR = 120 Days

4.1.3 Calmar Ratio

Let R_A(T) = average annual return over [0, T], and MDD(T) = average annual MDD over [0, T].

$$Calmar(T) = \frac{R_A(T)}{MDD(T)} \tag{4.2}$$

The Calmar ratio is a risk adjusted measure which resembles the Sharpe ratio. Both ratios are constructed by dividing return by risk. In the case of the Sharpe ratio the risk factor is the standard deviation of the returns. For the Calmar ratio, the risk factor is the MDD of the cumulative returns. Suprisingly, the MDD Constraint method generally had a lower Calmar ratio than the benchmarks. However, the Min-MDD algorithm produced the highest Calmar ratios amongst all of the methods tested and generally beat the benchmarks.

4.1.4 Clustering

Clustering did not appear to improve the performance of any of the tested algorithms. In several cases it diminished performance. Figure 4.5 shows the deterioration of the Calmar ratio for the Min-MDD algorithm when clustering is used. This is not the only example of performance degradation, but it is an important example as the Min-MDD algorithm without clustering produced some of the highest Calmar ratios. Thus clustering harmed the primary reason for using the Min-MDD allocation algorithm.

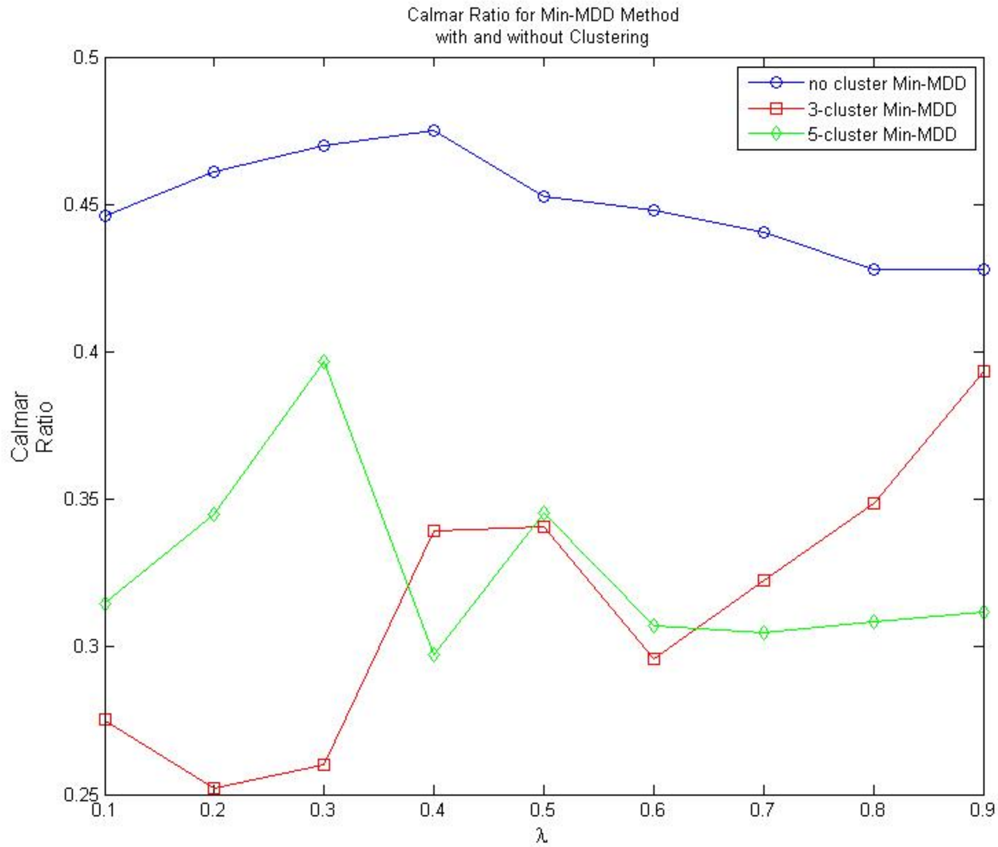


Figure 4.5: Calmar Ratio for Clustered and Unclustered Min-MDD

4.2 Min-MDD During Growth Periods and Recessions

Min-MDD was the only optimization algorithm that produced results that beat the benchmarks in all three measures of performance. To better understand why it was able to do so, four additional simulations were conducted using the benchmark algorithms and the Min-MDD algorithm. Each simulation uses data from a different time period. Two simulations use data from growth periods, periods of continuous stock price growth. The other two use data from recessions, periods of continuous decreases in stock price. The performance metrics collected from each simulation can be found on the following page.

Table 4.1: Growth Period (07/1/1996–06/16/2000) Performance

	Volatility Weighted	Uniform	MDD-Minimization
Sharpe Ratio	1.5334	1.5009	1.3117
Rate of Return	0.1589	0.1682	0.1566
Calmar Ratio	2.7391	2.7165	2.4085

Table 4.2: Growth Period (05/1/2003–04/23/2007) Performance

	Volatility Weighted	Uniform	MDD-Minimization
Sharpe Ratio	0.9742	1.0405	0.984
Rate of Return	0.1525	0.1688	0.1701
Calmar Ratio	1.3976	1.4588	1.6064

Table 4.3: Recession (09/11/2000–04/02/2003) Performance

	Volatility Weighted	Uniform	MDD-Minimization
Sharpe Ratio	-0.5635	-0.6072	-0.4387
Rate of Return	-0.1214	-0.1444	-0.804
Calmar Ratio	-1.9208	-2.2932	-1.1842

Table 4.4: Recession (05/1/2007–05/22/2009) Performance

	Volatility Weighted	Uniform	MDD-Minimization
Sharpe Ratio	-0.5096	-0.4896	-0.3276
Rate of Return	-0.1616	-0.1754	-0.0752
Calmar Ratio	-2.1804	-2.2479	-1.1337

5. DISCUSSION AND CONCLUSION

5.1 Mean-Var and MDD Constraint Methods

These methods are undesirable for the same reasons. Whenever they produced large Sharpe ratios, they also produced large drawdowns and low Calmer ratios. On the occasion that they produced large Calmar ratios, they also produced low Sharpe ratios. As such, an investor who is aware of each type of risk would never chose to use one of these methods.

5.2 Min-MDD Method

The results of the Min-MDD algorithm show that MDD should be considered seriously as a measure of risk. The additional simulations to test Min-MDD's performance in different market conditions provide solid evidence as to why MDD is a useful measure of risk. Min-MDD performed far better during the recession periods, especially during the 2007–2009 period, than the benchmark algorithms. An analytical study of MDD by Magdon-Ismail and Atiya found that instruments with negative returns can actually be beneficial to a portfolio when considering the Calmar ratio as an acceptable risk adjusted measure of performance (Magdon-Ismail & Atiya, 2004). These impressive results during recessions come without sacrificing performance during boom periods, as the Min-MDD algorithm performed similarly to the benchmarks in both growth periods. These results make sense intuitively as minimizing the MDD should be expected to protect an investment from large losses.

5.3 Clustering

A possible explanation for why clustering fails to improve or maintain performance is that not enough clusters are being used. Having more clusters could improve the optimality of the allocation amongst meta-stocks as the presence of more assets generally allows for greater diversification. This could then improve the optimality of overall allocation. The downside of increasing the number of clusters is that the size of each cluster will decrease potentially ruining the allocations withing clusters. Further research should be conducted to determine the optimal number of clusters.

5.4 CONCLUSION

The performances of various portfolio allocation algorithms were compared in a simulation environment using real stock market data. The traditional Markowitz Mean-Var method and a new MDD Constraint method were found to sometimes outperform benchmarks in one or two performance metrics, but never in all three tested metrics. These methods more often underperformed the benchmarks. The Sharpe ratio was also shown to be insufficient as a measure of risk when investors are concerned about large drawdowns. Several configurations tested in simulation produced large Sharpe ratios and small Calmar ratios. It is the opinion of the author that MDD as a measure of risk should continue to be studied. Investors are already allocating in such a way as to prevent large drawdowns, and the Min-MDD optimization algorithm has been shown in this research to produce good performance results. Specifically, Min-MDD with $phor$ = 500 days and $\lambda = 0.6$ or 0.7 which were found to beat the benchmarks in every performance metric considered. Table 5.1 shows the author's suggested choices for λ and $phor$ for each method when considering each measure of performance as important.

Table 5.1: Optimal λ and $phor$ for Different Risk Measures

Method	λ^*	$phor^*$	Annual Rate of Return	Sharpe ratio	Calmar ratio
Mean-Var	0.8	500	0.1027	0.6130	0.4106
Min-MDD	0.7	500	0.1155	0.7368	0.4405
MDD Constraint	0.6	500	0.1074	0.5989	0.4055

REFERENCES

[1] Burghardt G., Duncan R., and Lianyan L. (2003). Understanding Drawdowns. *Carr Futures*, working paper.

[2] Cheklov, Alexei, Uryasev, Stanislav P. and Zabarankin M. (2005). Portfolio Optimization with Drawdown Constraints. *International Journal of Theoretical and Applied Finance*, *8*(1), 13-58.

[3] Defusco R., Karels G., and Muralidhar K. (1996). Skewness Persistence in US Common Stock Returns: Results from Bootstrapping Tests. *Journal of Business Finance & Accounting*, *23*(8), 1183-1195.

[4] Hamelink F., and Hoesli M. (2004). Maximum drawdown and the allocation to real estate. *Journal of Property Research*, *21*(1), 5-29.

[5] Jondeau E., and Rockinger M. (2006). Optimal Portfolio Allocation under Higher Moments. *European Financial Management*, *12*(1), 29-55.

[6] Magdon-Ismail M., and Atiya A. (2004). An Analysis of the Maximum Drawdown Risk Measure. *Risk Magazine*, *17*(10), 99-102.

[7] Markowitz H. (1999). The Early History of Portfolio Theory: 1600-1960. *Financial Analysis Journal*, *55*(5), 5-16.

[8] Markowitz H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77-91.

[9] Menezes C., Hanson D. (1970). On the Theory of Risk Aversion. *International Economic Review*, *11*(3), 481-487.

[10] Pratt J. (1964). Risk Aversion in the Small and Large,*The Econometric Society*, *32*(1/2), 122-136.

[11] Simkowitz M., and Beedles W (1978). Diversification in a Three-Moment World. *The Journal of Financial and Quantitative Analysis*, *13*(5), 927-941.

[12] Singleton, J.C., and Wingender J (1986). Skewness Persistence in Common Stock Returns. *The Journal of Financial and Quantitative Analysis*, *21*(3), 335-341.

[13] Tobin, J. (1958). Liquidity preference as behavior towards risk, *Review of Economic Studies*, *25*, 65-86.

APPENDIX A
LINEAR PROGRAMS IN STANDARD FORM

The linear programs used in this research are presented below in standard form.

$$
\begin{aligned}
&\min f' \cdot x,\\
&s.t.\\
&A \cdot x \leq b,\\
&A_{eq} \cdot x \leq b_{eq},\\
&lb \leq x \leq ub
\end{aligned}
$$

A.1 Maximizing Return Subject to an MDD Constraint

Let I_T be a TxT identity matrix, 0_T be a Tx1 matrix of zeros, 0_n be a Nx1 matrix of zeros, and R_T be the vector containing the cumulative returns for each stock at the end of the [0,T] period.

$$
x = \begin{bmatrix} z_0 \\ z \\ \Theta \end{bmatrix}, \quad f = -\begin{bmatrix} 0 \\ 0_T \\ R_T' \end{bmatrix}
$$

$$
A = \begin{bmatrix} 0_T & I_T & -R \\ 0_T & -I_T & R \\ \begin{bmatrix} I_T & 0_T \end{bmatrix} + & \begin{bmatrix} 0_T & -I_T \end{bmatrix} & 0_T \cdot 0_n' \end{bmatrix}, \quad b = \begin{bmatrix} \lambda \cdot MDD(Uni) \cdot 1_T \\ 0_T \\ 0_T \end{bmatrix}
$$

$$
A_{eq} = \begin{bmatrix} 1 & 0_T' & 0_n' \\ 0 & 0_T' & 1_n' \end{bmatrix}, \quad b_{eq} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}
$$

$$lb = \begin{bmatrix} 0 \\ 0_T \\ 0_n \end{bmatrix}, \quad ub = \begin{bmatrix} \infty \\ \infty \cdot 1_T \\ 1 \end{bmatrix}$$

A.2 Minimizing MDD Subject to a Return Constraint

Let I_T be a TxT identity matrix, 0_T be a Tx1 matrix of zeros, 0_n be a Nx1 matrix of zeros, and R_T be the vector containing the cumulative returns for each stock at the end of the [0,T] period. $Target$ specifies the level of return for which risk is being minimized.

$$x = \begin{bmatrix} z_0 \\ z \\ \Theta \\ m \end{bmatrix}, \quad f = \begin{bmatrix} 0 \\ 0_T \\ 0_n \\ 1 \end{bmatrix}$$

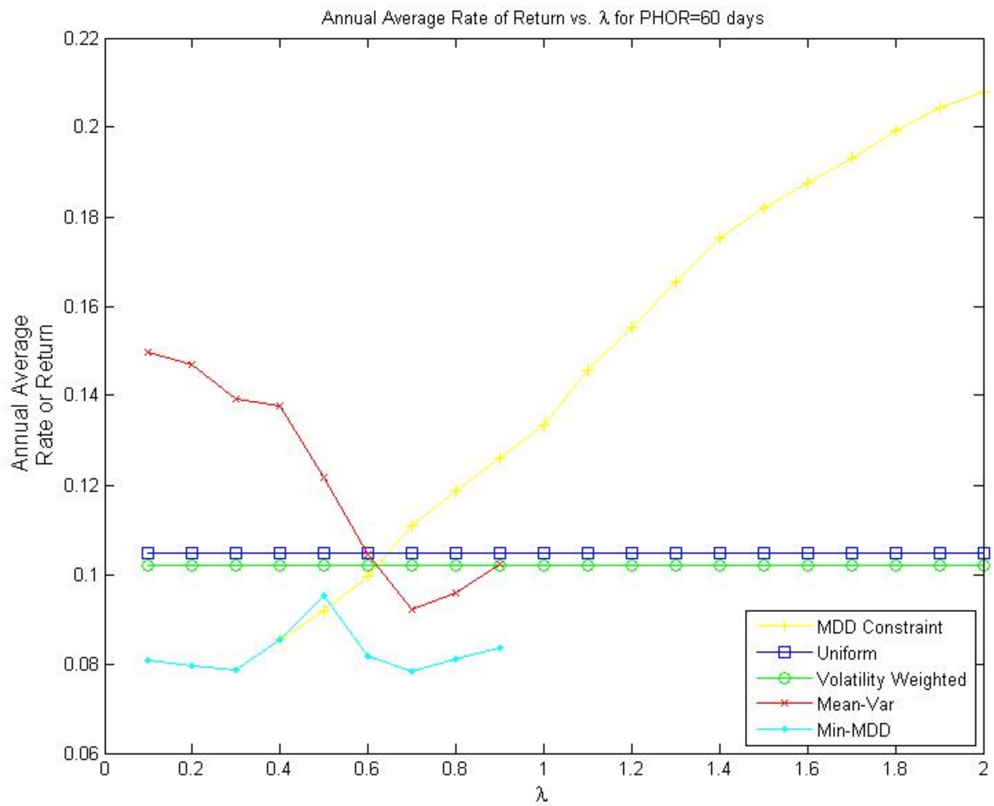
$$A = \begin{bmatrix} 0 & 0_T^{'} & -R_T & 0 \\ 0_T & I_T & -R & -1_T \\ 0_T & -I_T & R & 0_T \\ \begin{bmatrix} I_T & 0_T \end{bmatrix} + & \begin{bmatrix} 0_T & -I_T \end{bmatrix} & 0_T \cdot 0_n^{'} & 0_T \end{bmatrix}, \quad b = \begin{bmatrix} -target \\ -1_T \\ 0_T \\ 0_T \end{bmatrix}$$

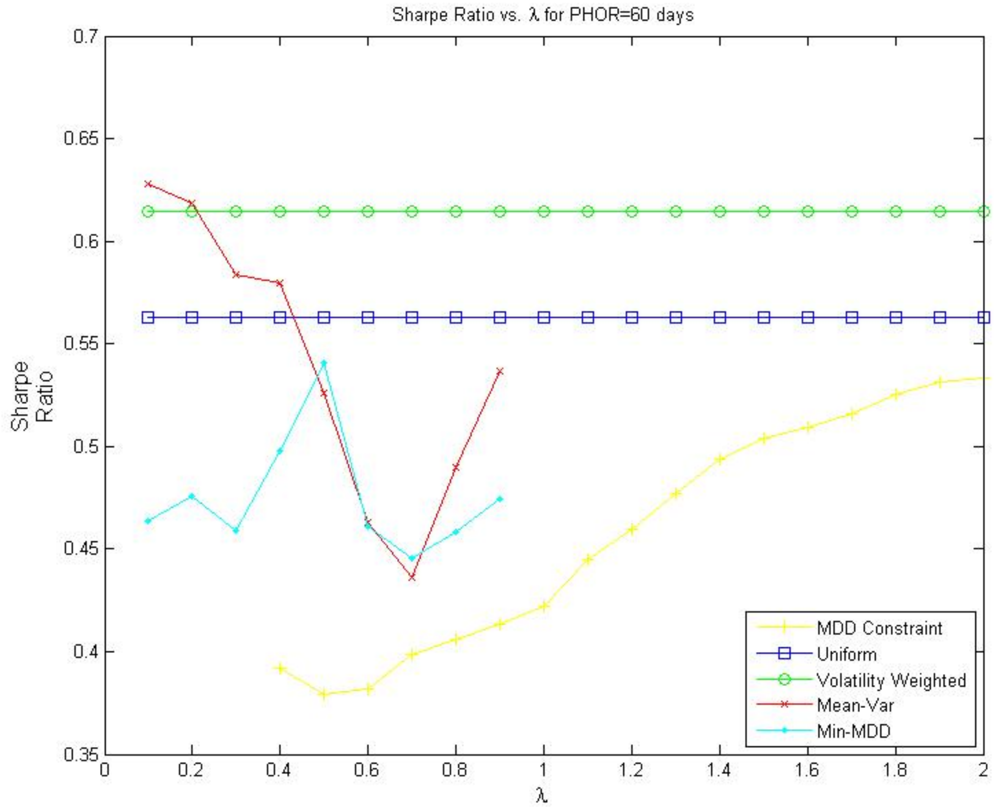
$$A_{eq} = \begin{bmatrix} 1 & 0_T^{'} & 0_n^{'} & 0 \\ 0 & 0_T^{'} & 1_n^{'} & 0 \end{bmatrix}, \quad b_{eq} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

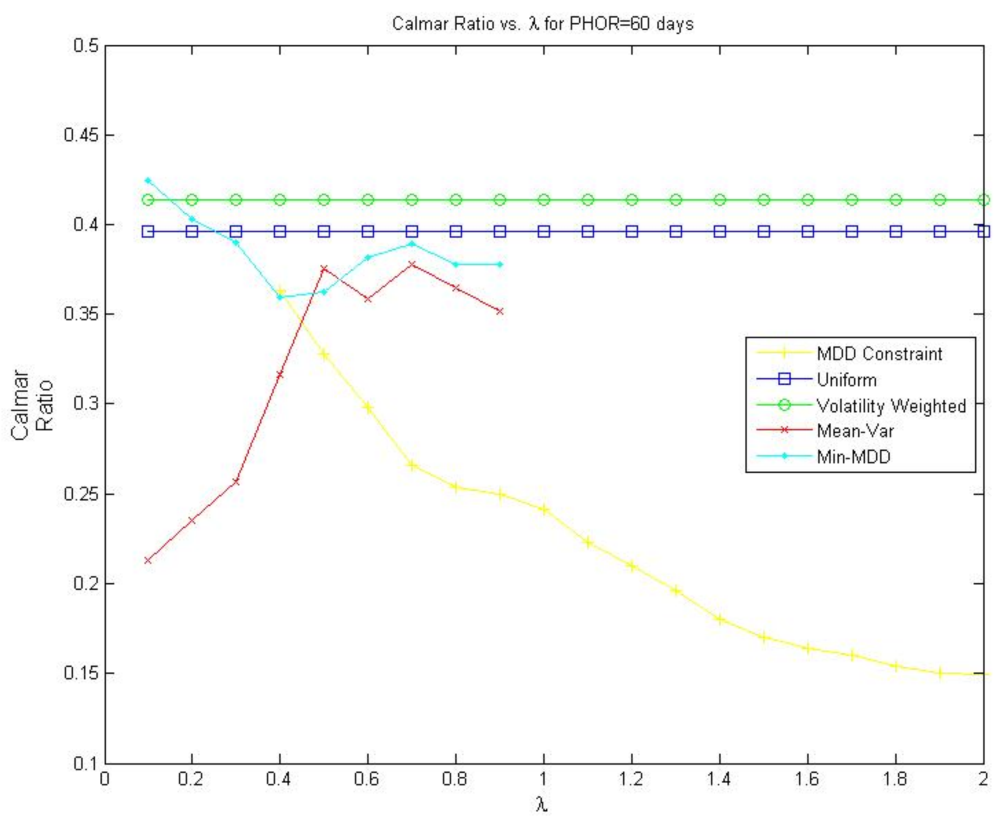
$$lb = \begin{bmatrix} 0 \\ 0_T \\ 0_n \\ 0 \end{bmatrix}, \quad ub = \begin{bmatrix} \infty \\ \infty \cdot 1_T \\ 1 \\ \infty \end{bmatrix}$$

APPENDIX B
ALL RESULTS

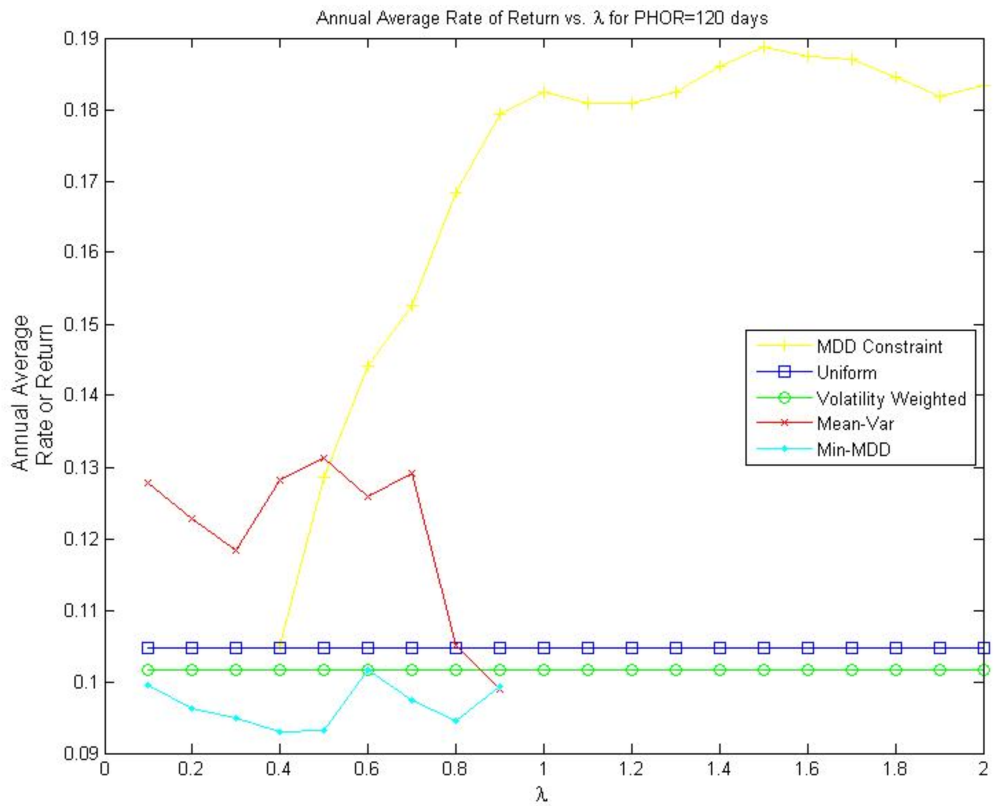
B.1 PHOR=60 days, No Clustering

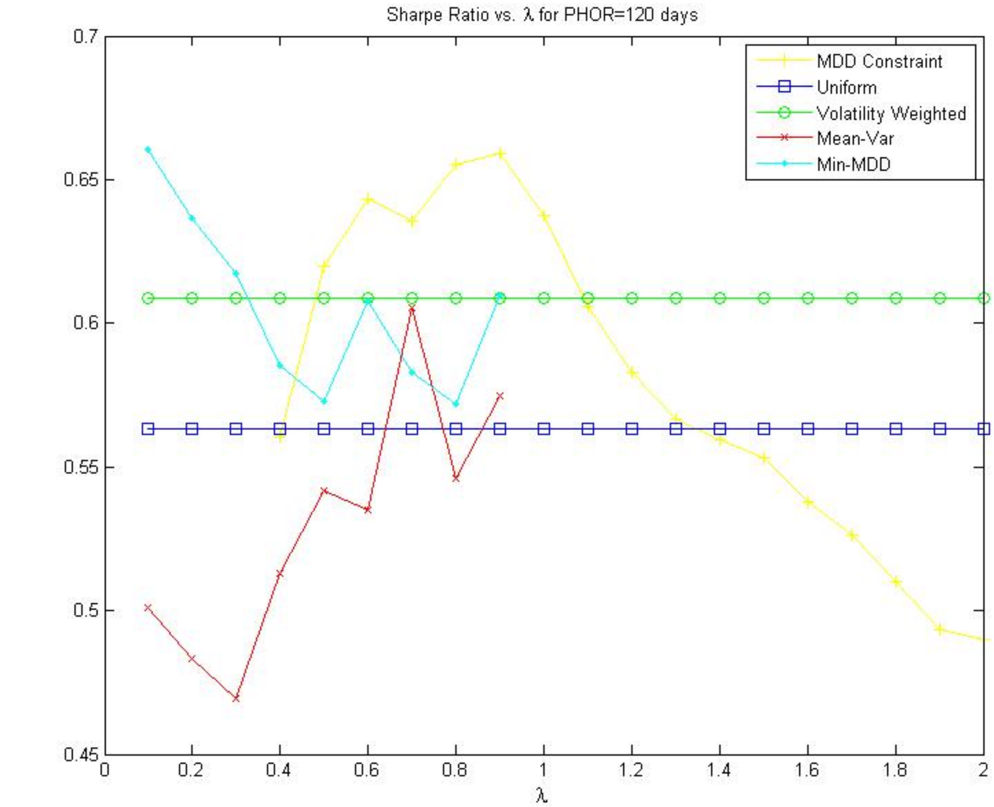


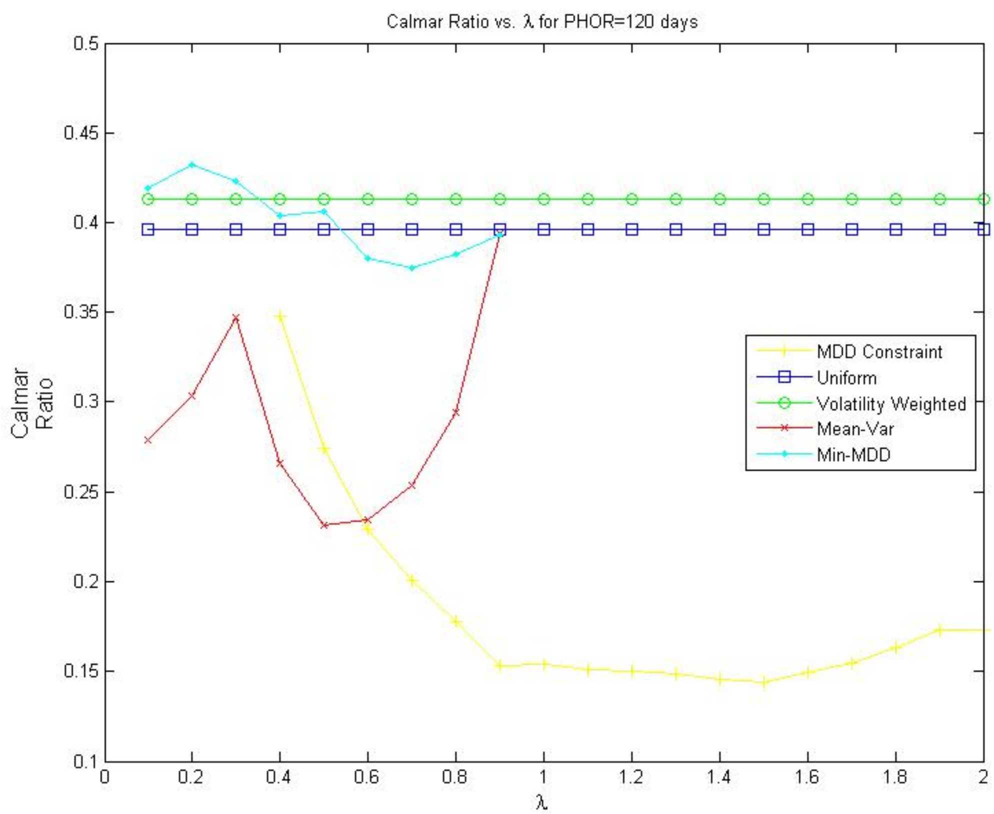

Sharpe Ratio vs. λ for PHOR=60 days
Sharpe Ratio
λ
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD


Calmar Ratio vs. λ for PHOR=60 days
Calmar Ratio
λ
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD

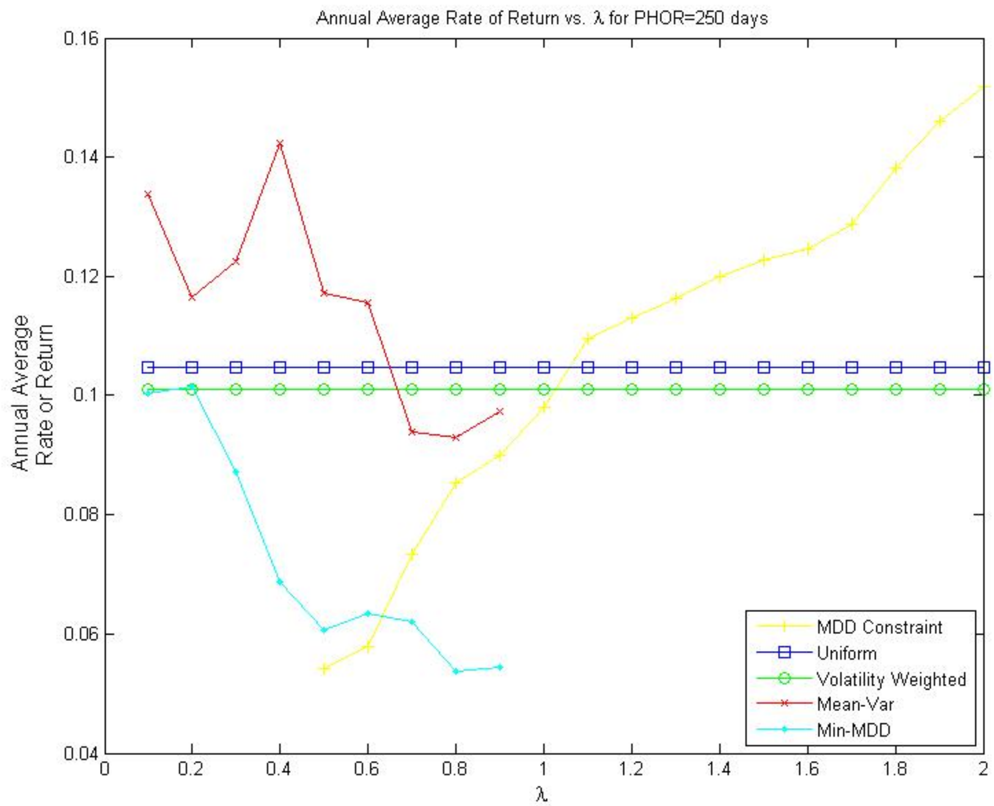
B.2 PHOR=120 days, No Clustering

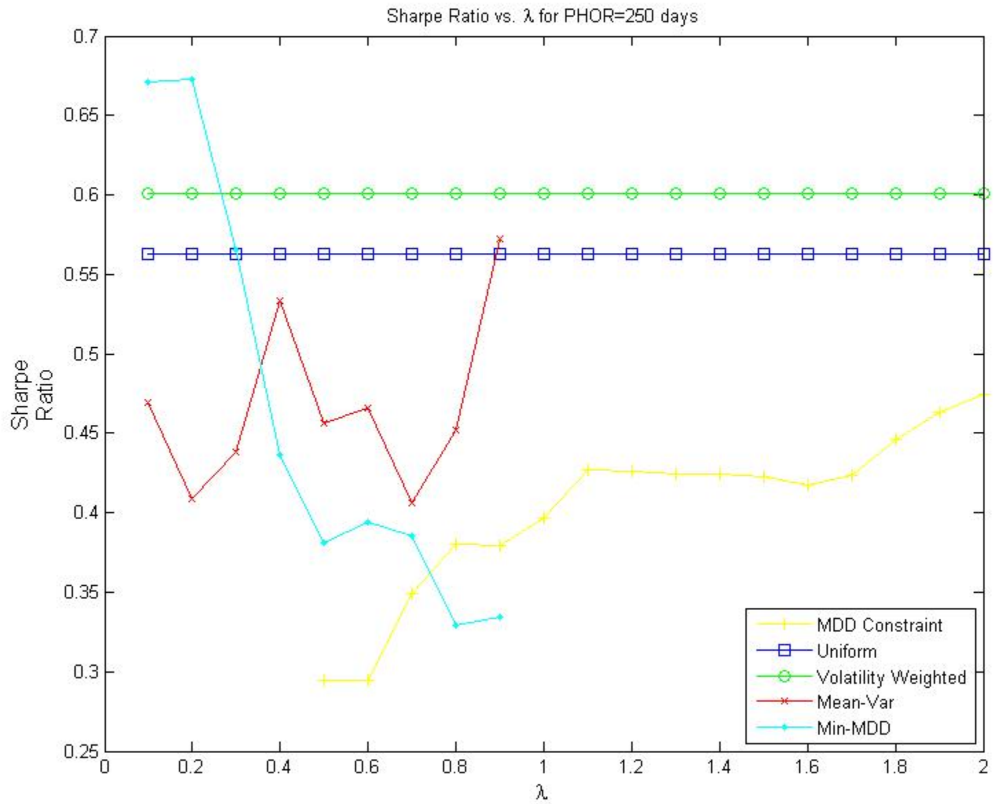



Sharpe Ratio vs. λ for PHOR=120 days
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD
λ

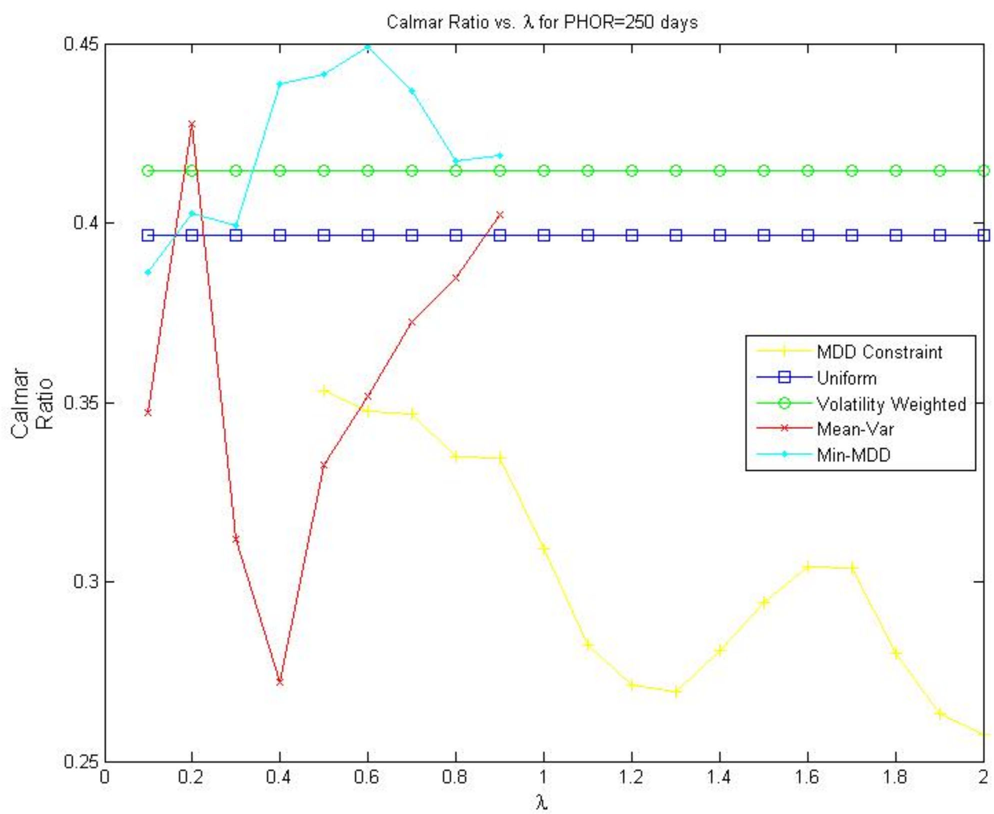

Calmar Ratio vs. λ for PHOR=120 days
Calmar Ratio
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD
λ

B.3 PHOR=250 days, No Clustering



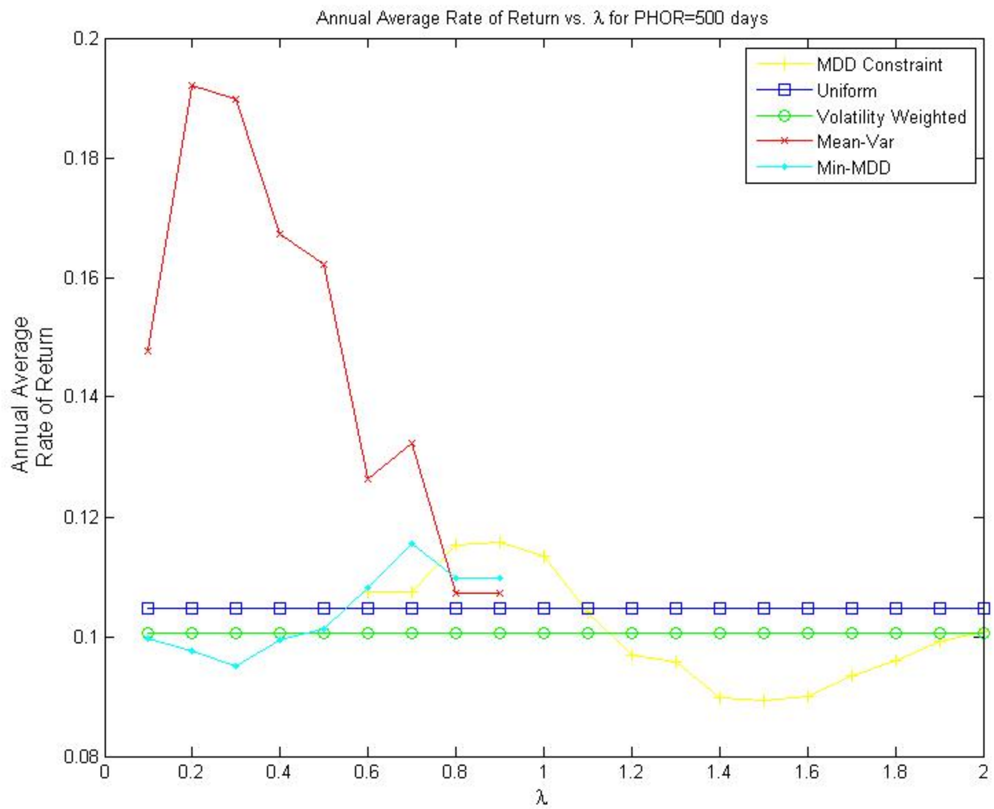


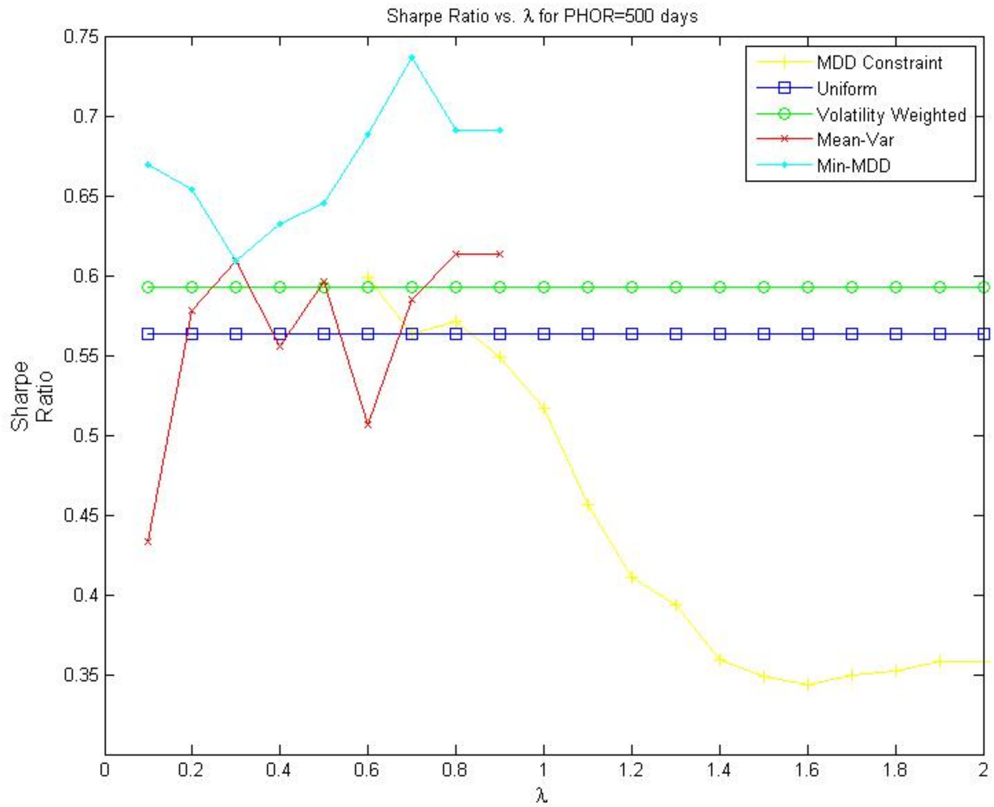
Sharpe Ratio vs. λ for PHOR=250 days
Sharpe Ratio
λ
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD

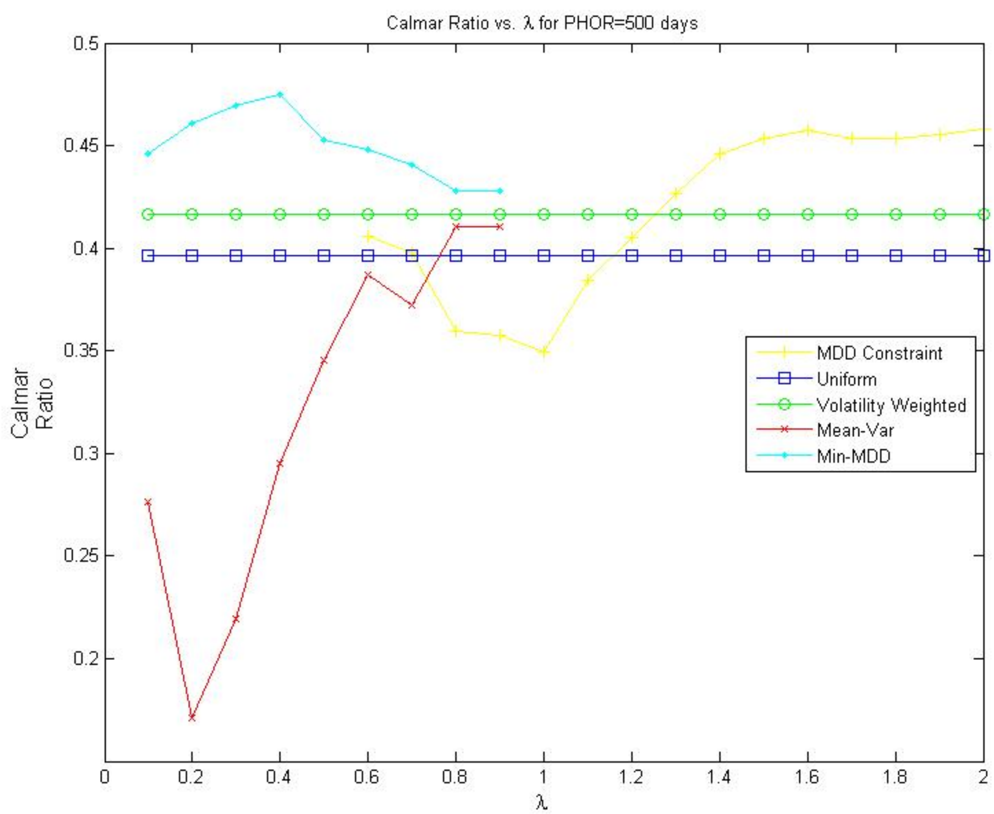


Calmar Ratio vs. λ for PHOR=250 days
Calmar Ratio
λ
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD

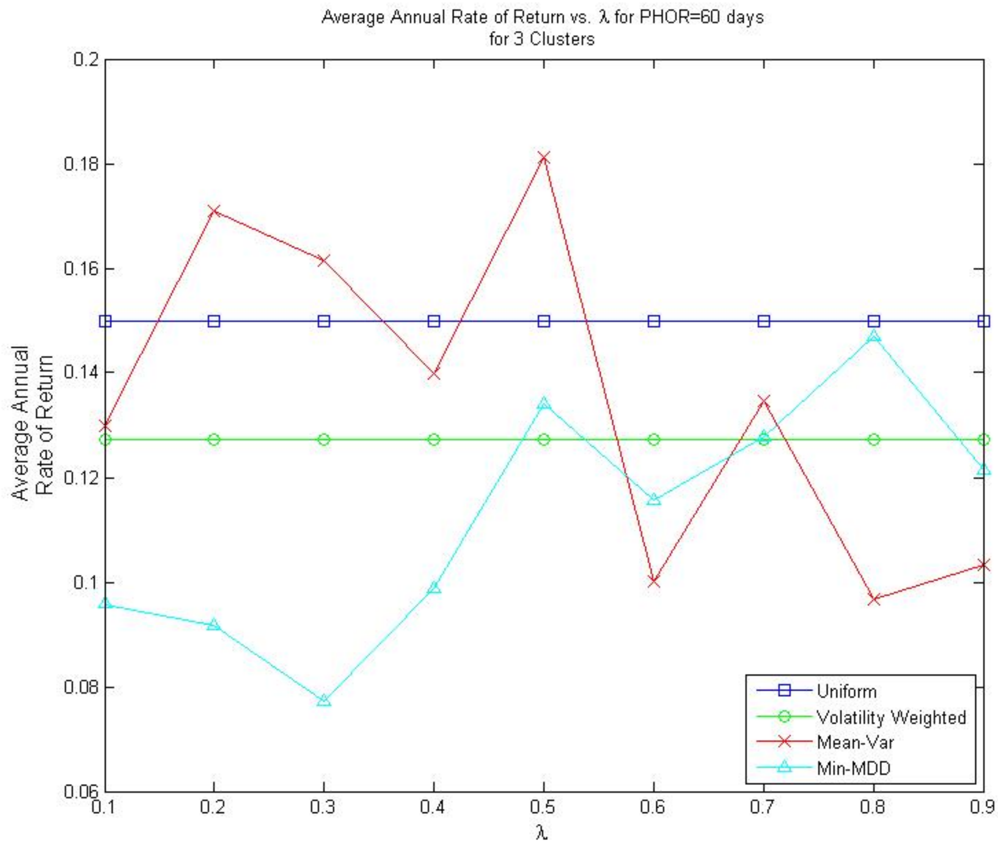
B.4 PHOR=500 days, No Clustering

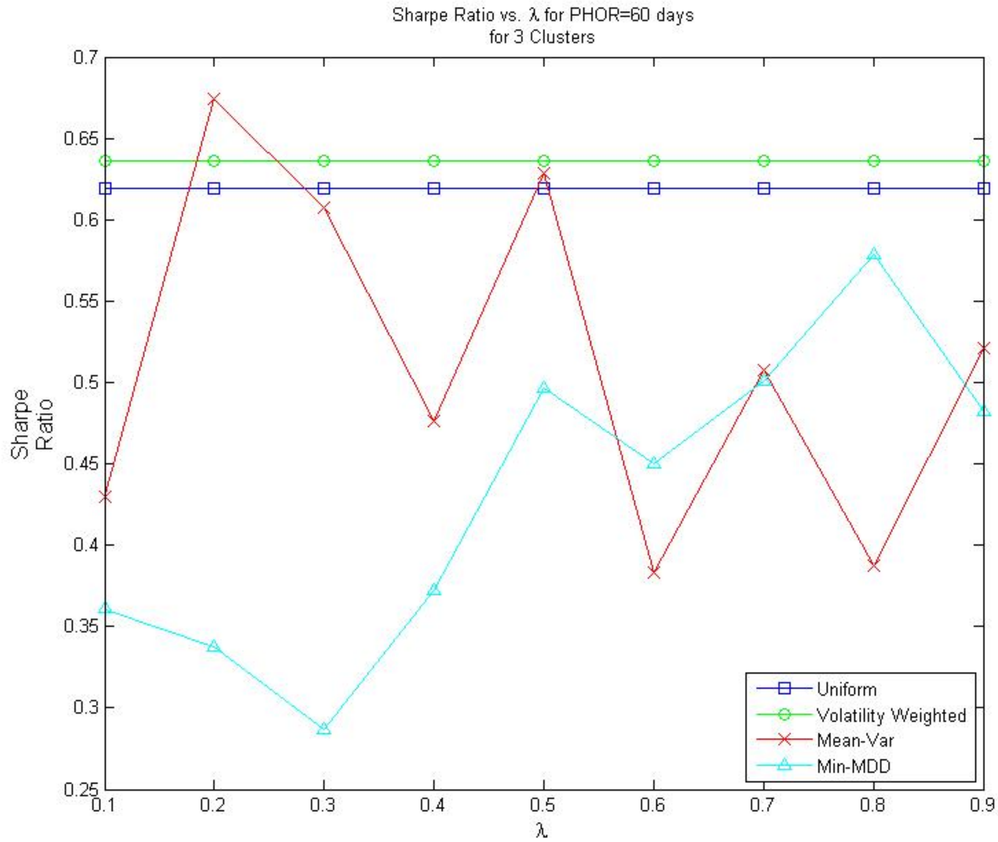


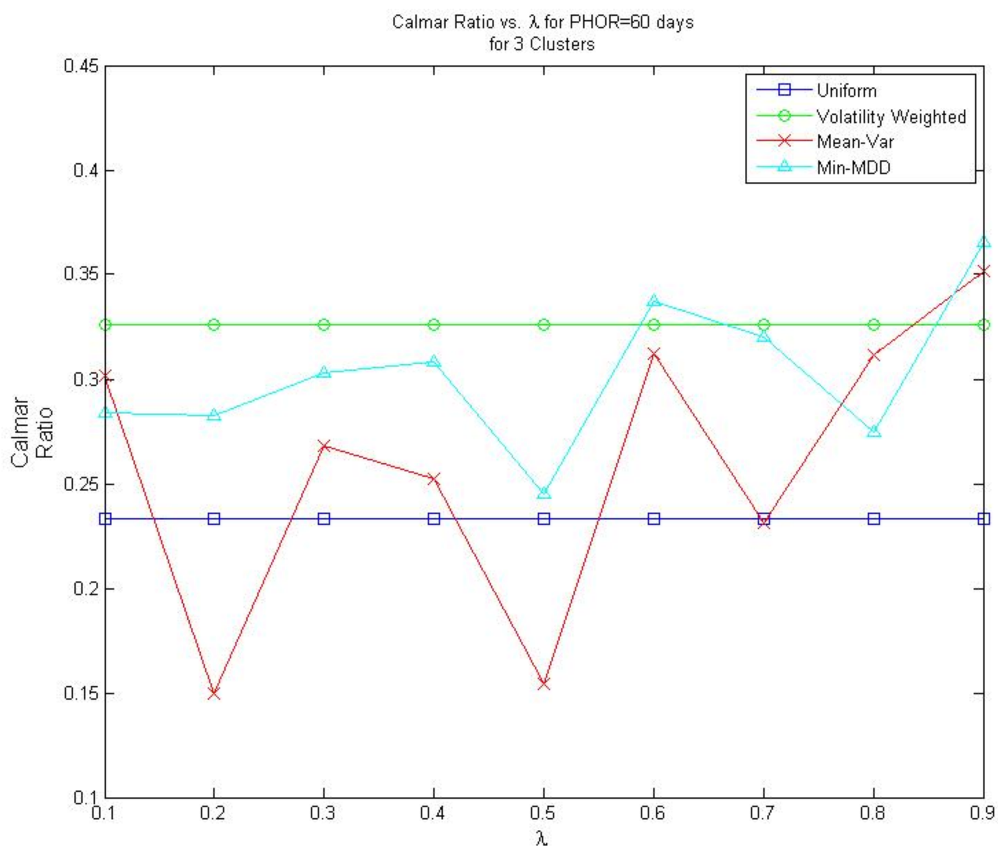

Sharpe Ratio vs. λ for PHOR=500 days
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD
Sharpe Ratio
λ
0.75
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2


Calmar Ratio vs. λ for PHOR=500 days
MDD Constraint
Uniform
Volatility Weighted
Mean-Var
Min-MDD
Calmar Ratio
λ
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2

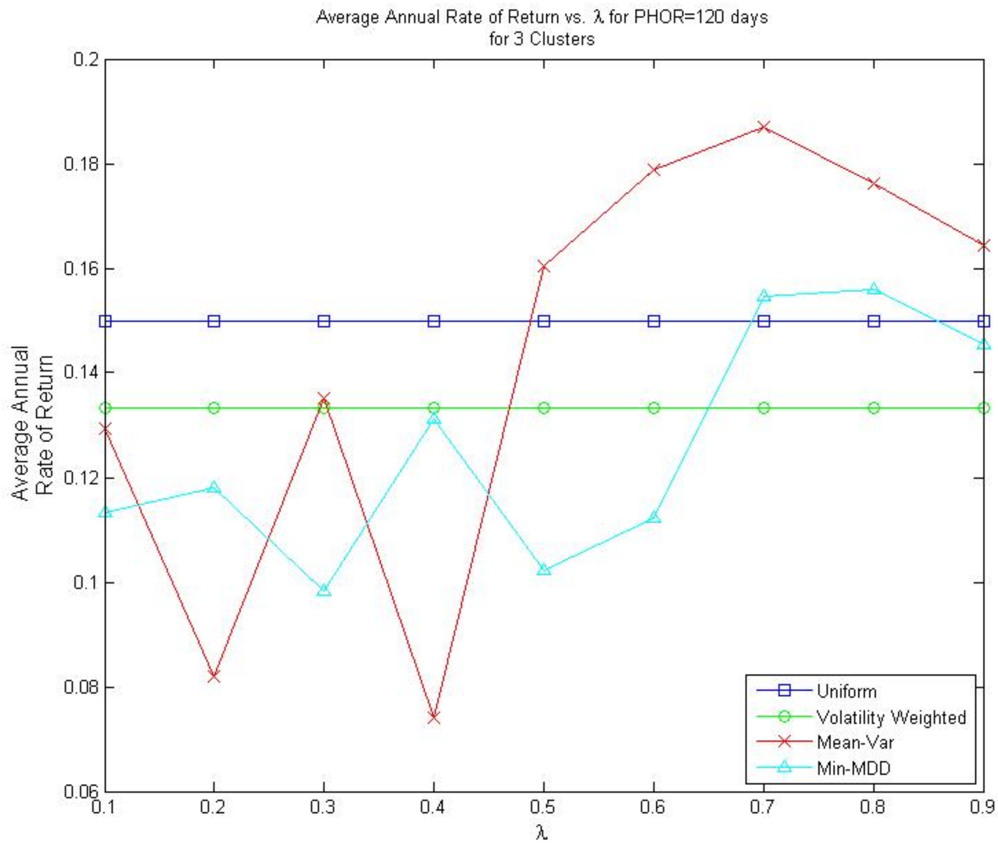
B.5 PHOR=60 days, 3-Clustering

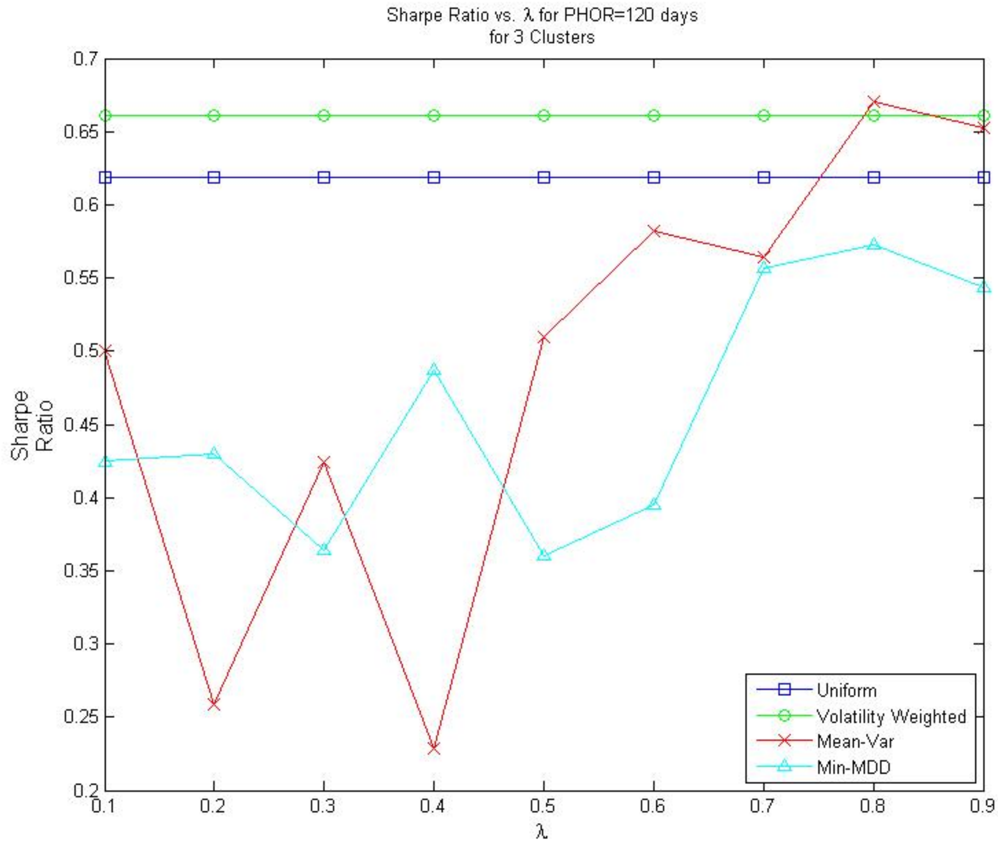



Sharpe Ratio vs. λ for PHOR=60 days
for 3 Clusters
Sharpe Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.3
0.25
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

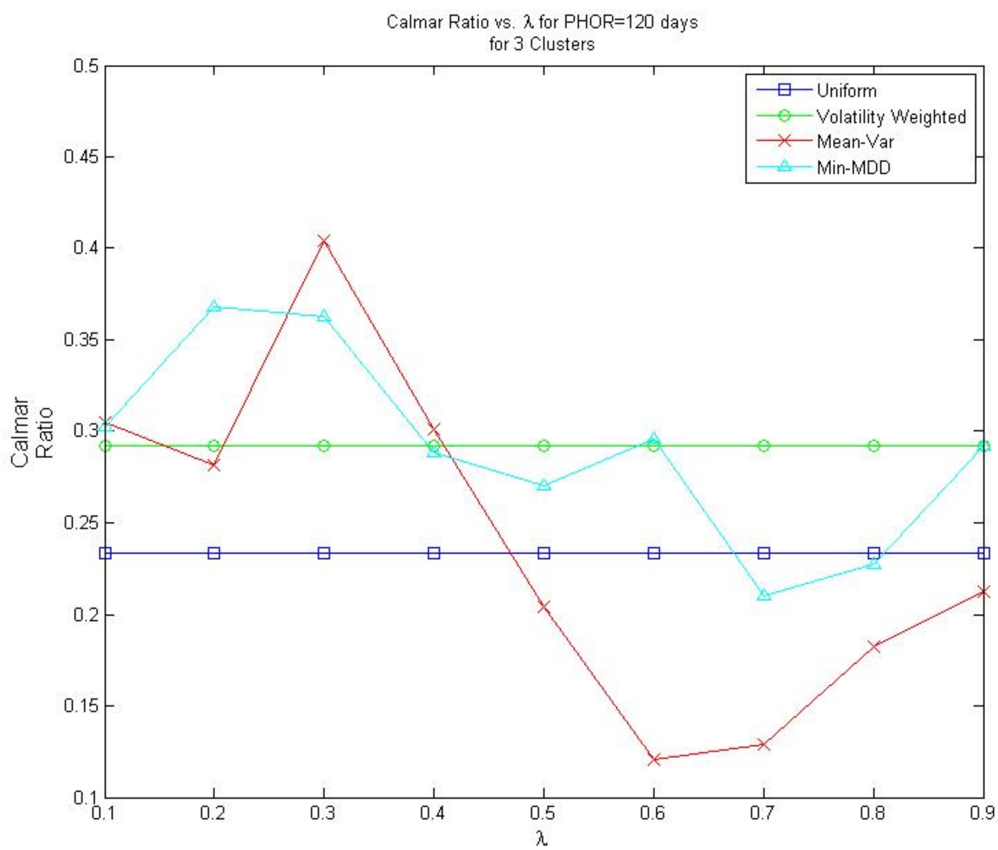

Calmar Ratio vs. λ for PHOR=60 days
for 3 Clusters
Calmar Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

B.6 PHOR=120 days, 3-Clustering



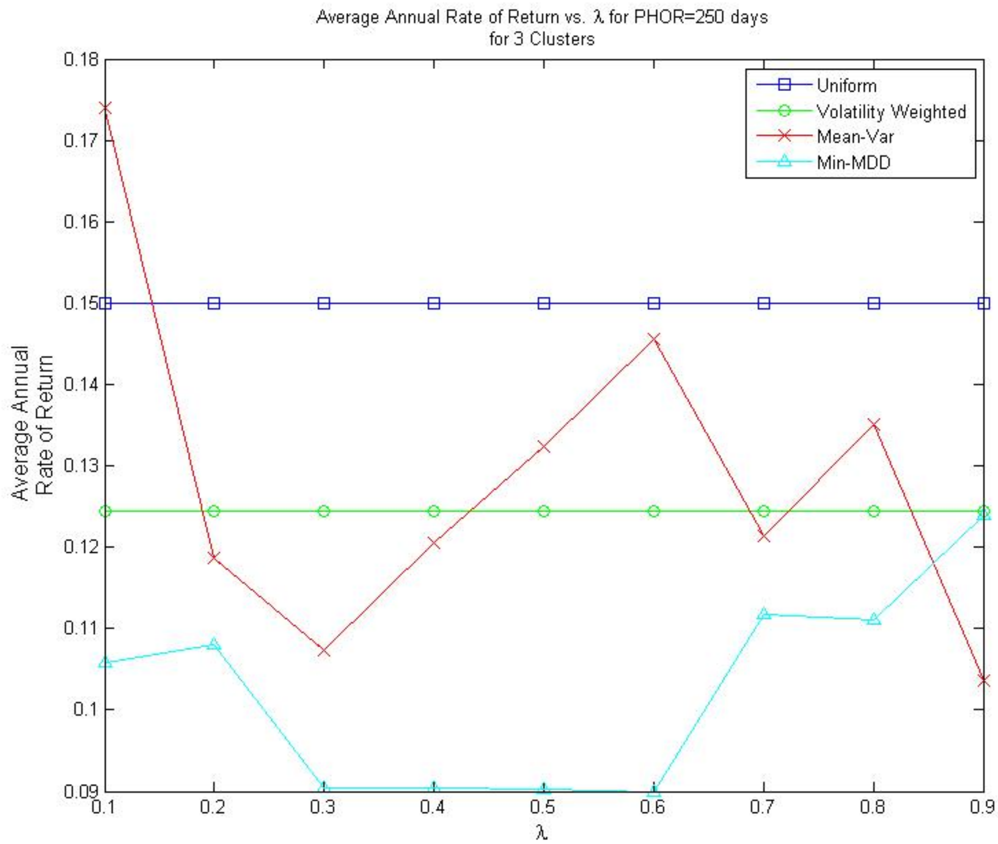


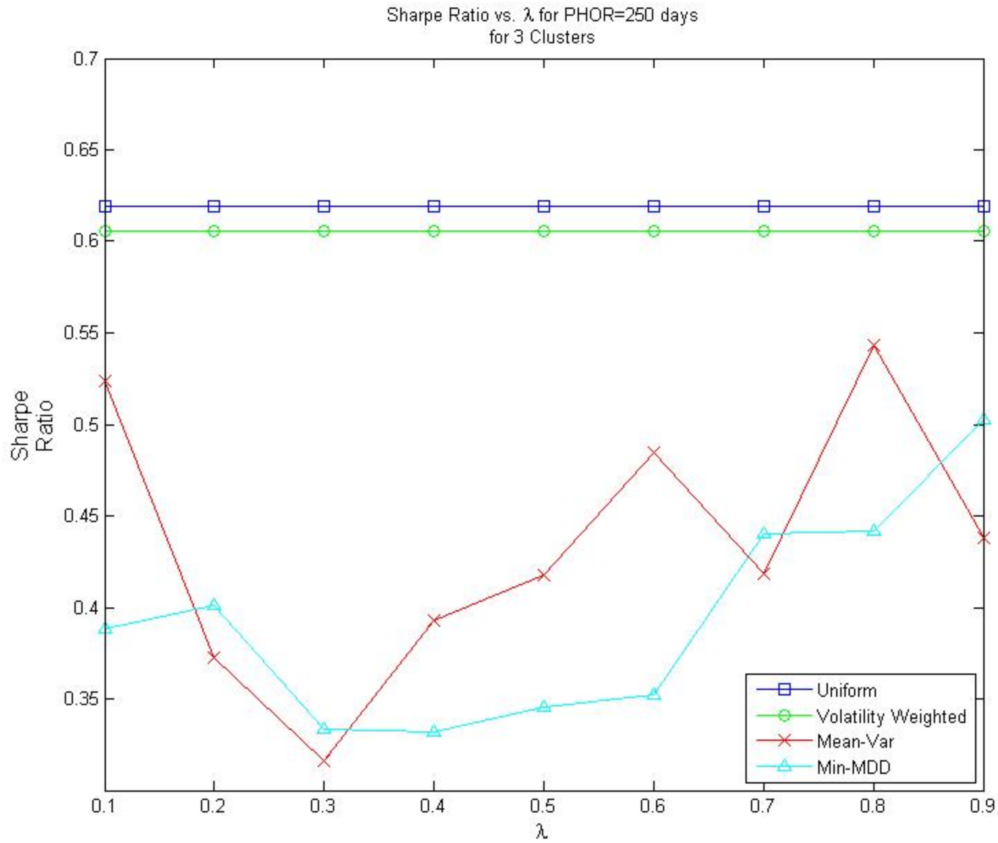
Sharpe Ratio vs. λ for PHOR=120 days
for 3 Clusters
Sharpe Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD

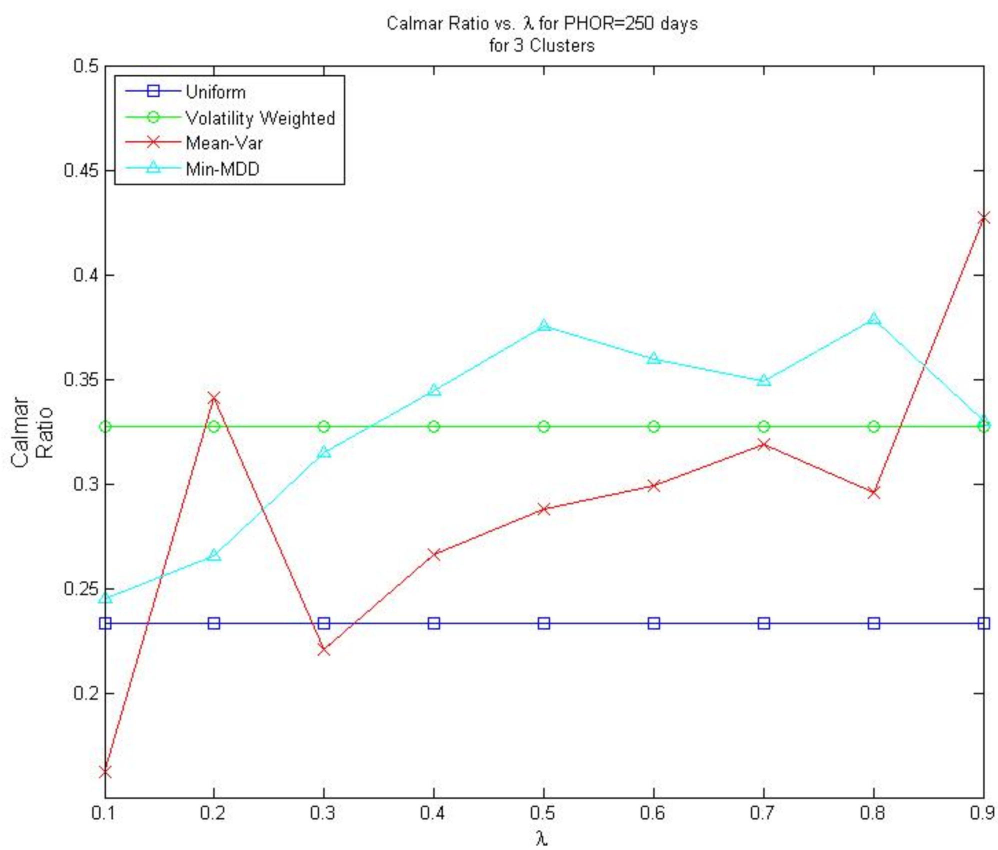


Calmar Ratio vs. λ for PHOR=120 days
for 3 Clusters
Calmar Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD

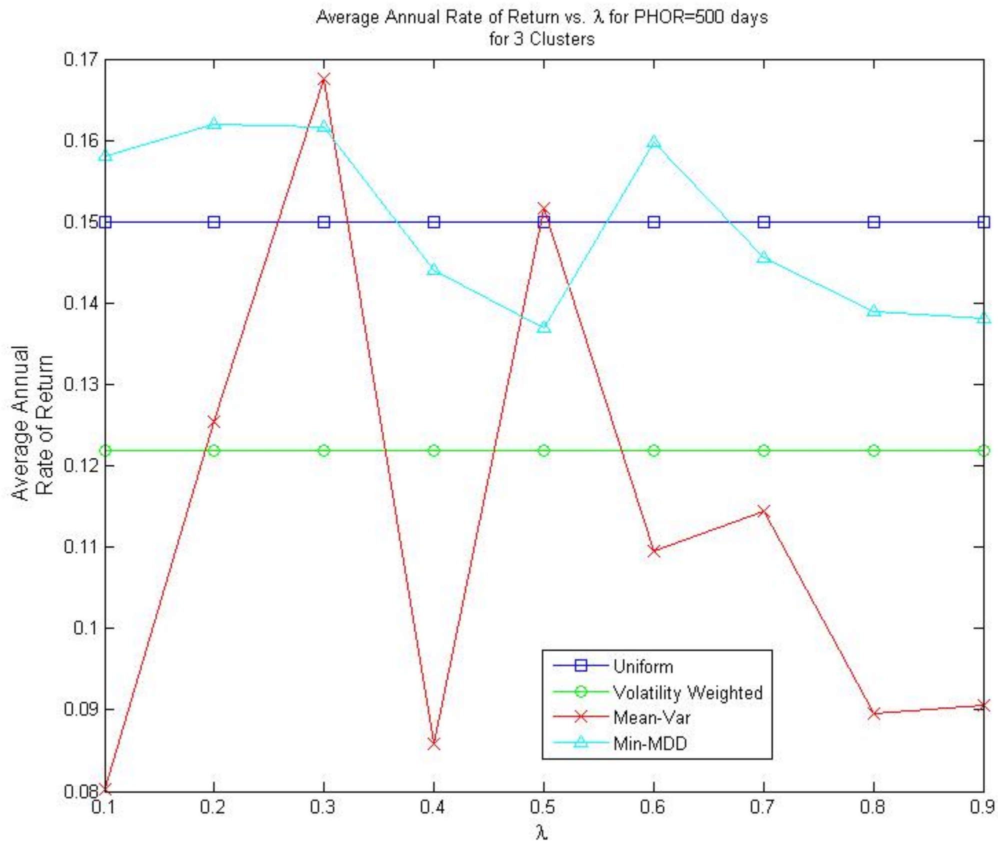
B.7 PHOR=250 days, 3-Clustering

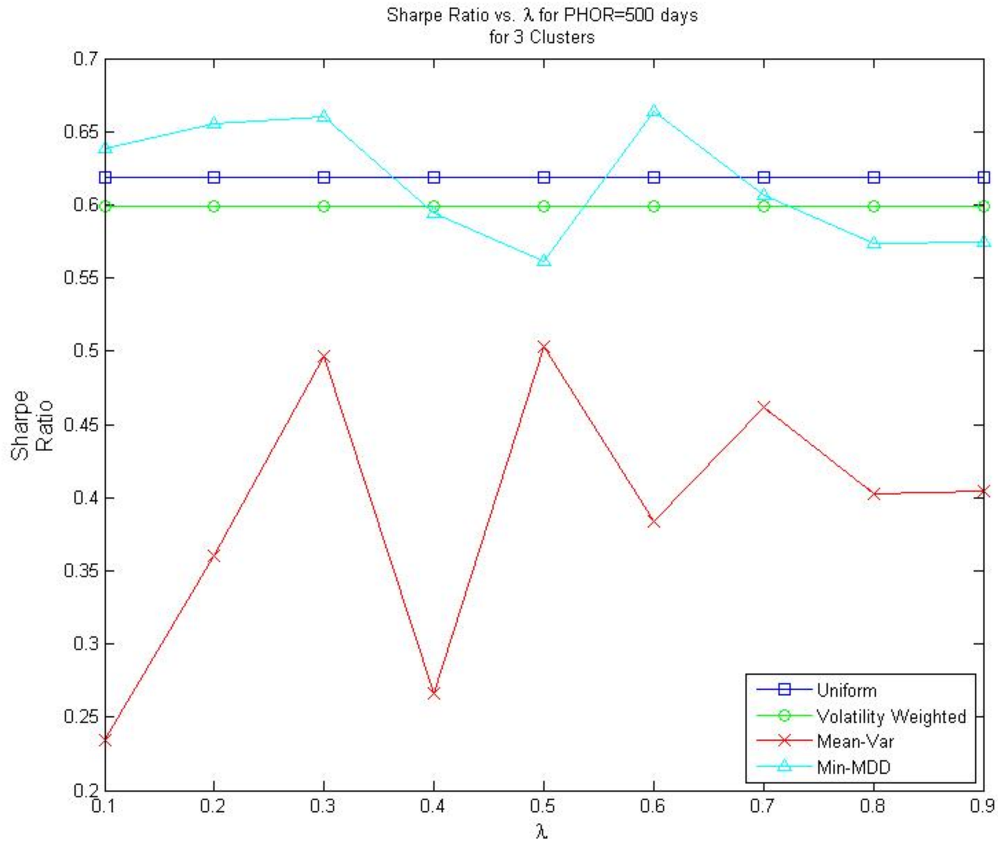


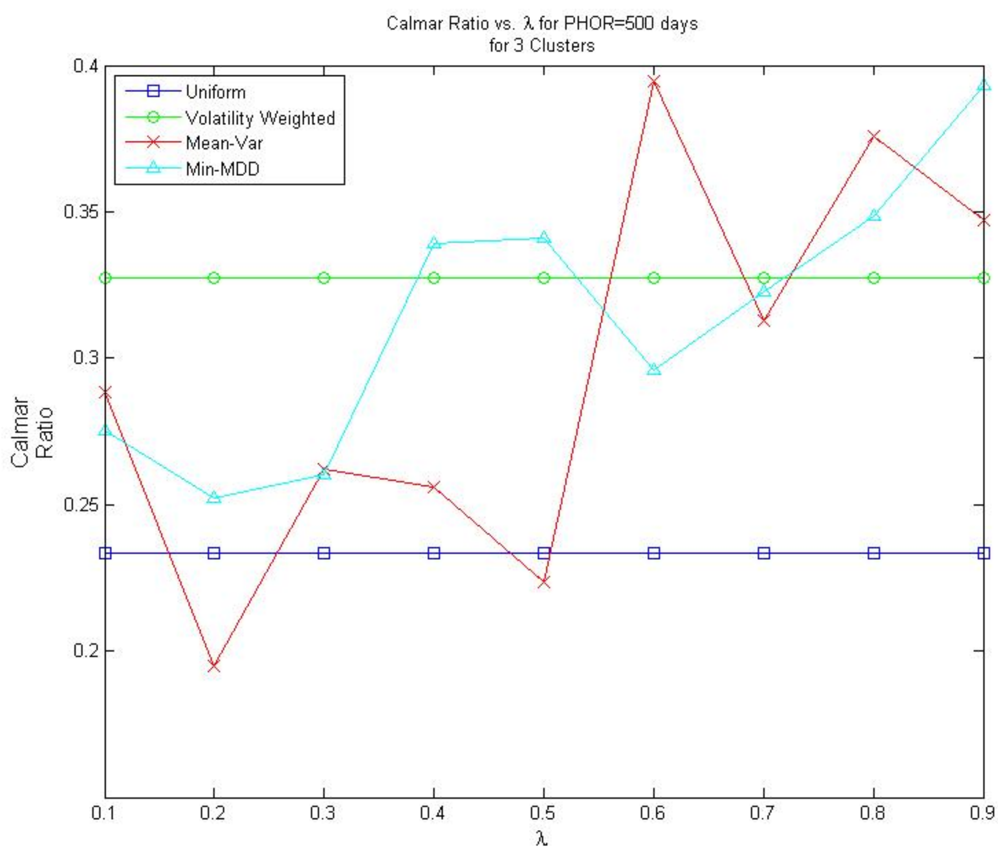

Sharpe Ratio vs. λ for PHOR=250 days
for 3 Clusters
Sharpe
Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.7
0.65
0.6
0.55
0.5
0.45
0.4
0.35
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Calmar Ratio vs. λ for PHOR=250 days
for 3 Clusters
Calmar
Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

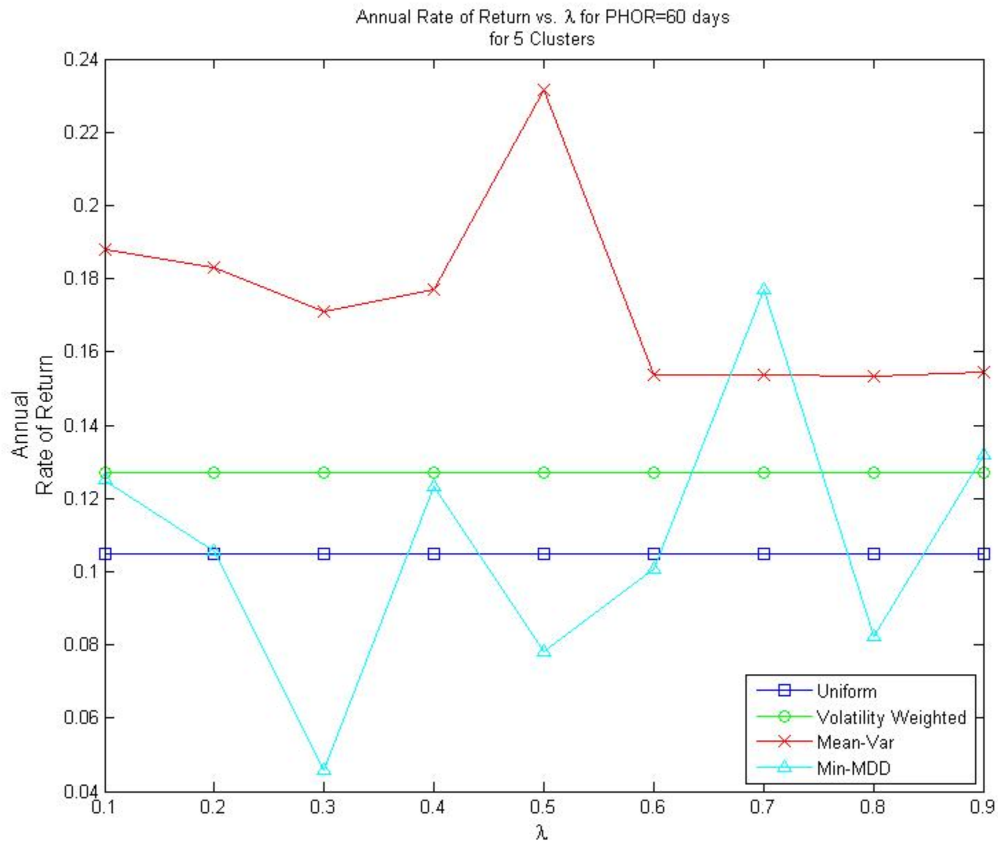
B.8 PHOR=500 days, 3-Clustering

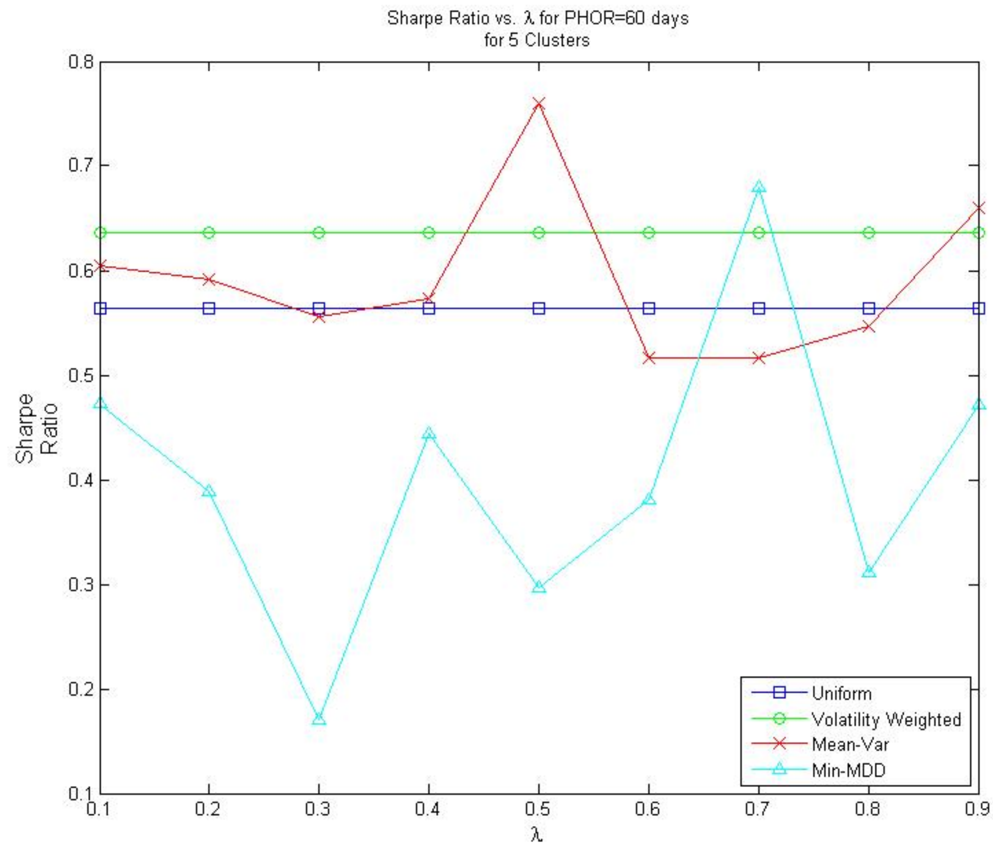


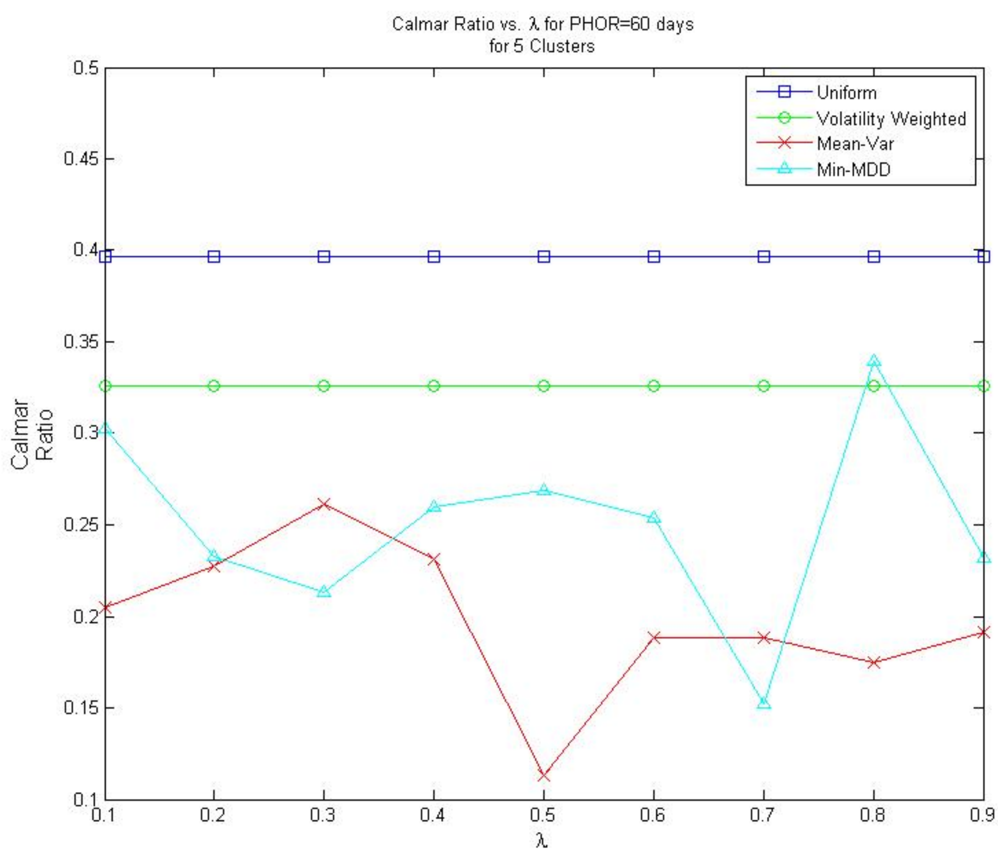

Sharpe Ratio vs. λ for PHOR=500 days
for 3 Clusters
Sharpe Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD


Calmar Ratio vs. λ for PHOR=500 days
for 3 Clusters
Calmar Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD

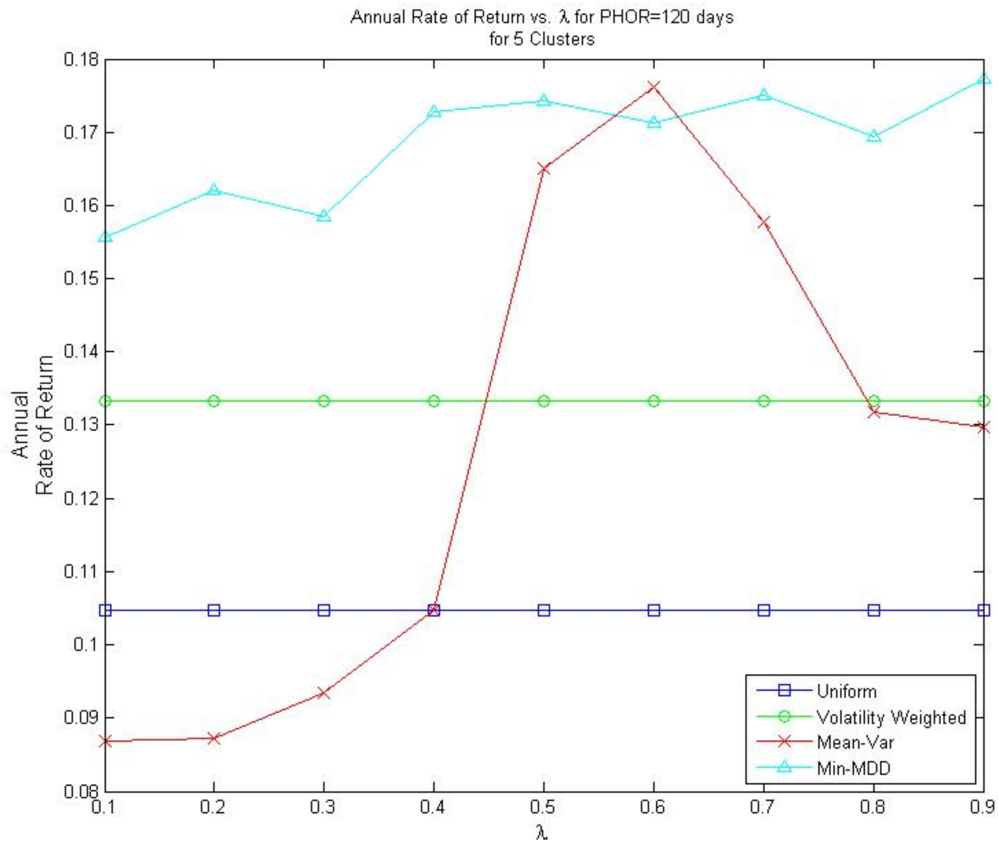
B.9 PHOR=60 days, 5-Clustering

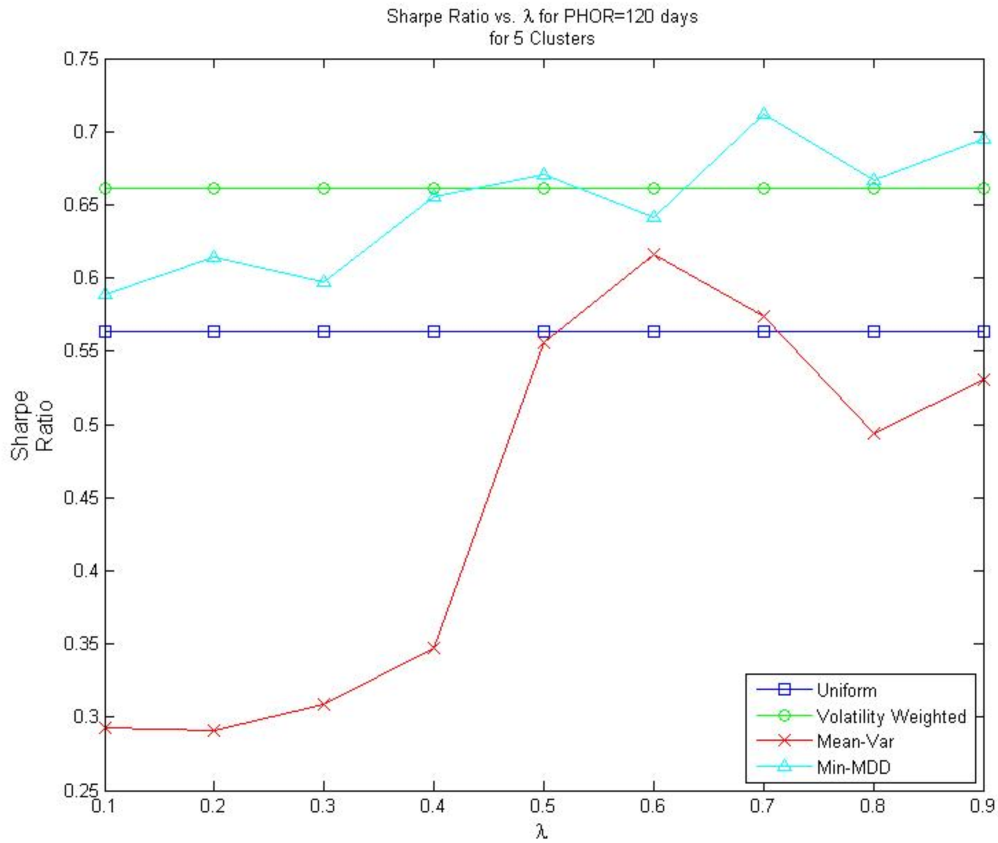


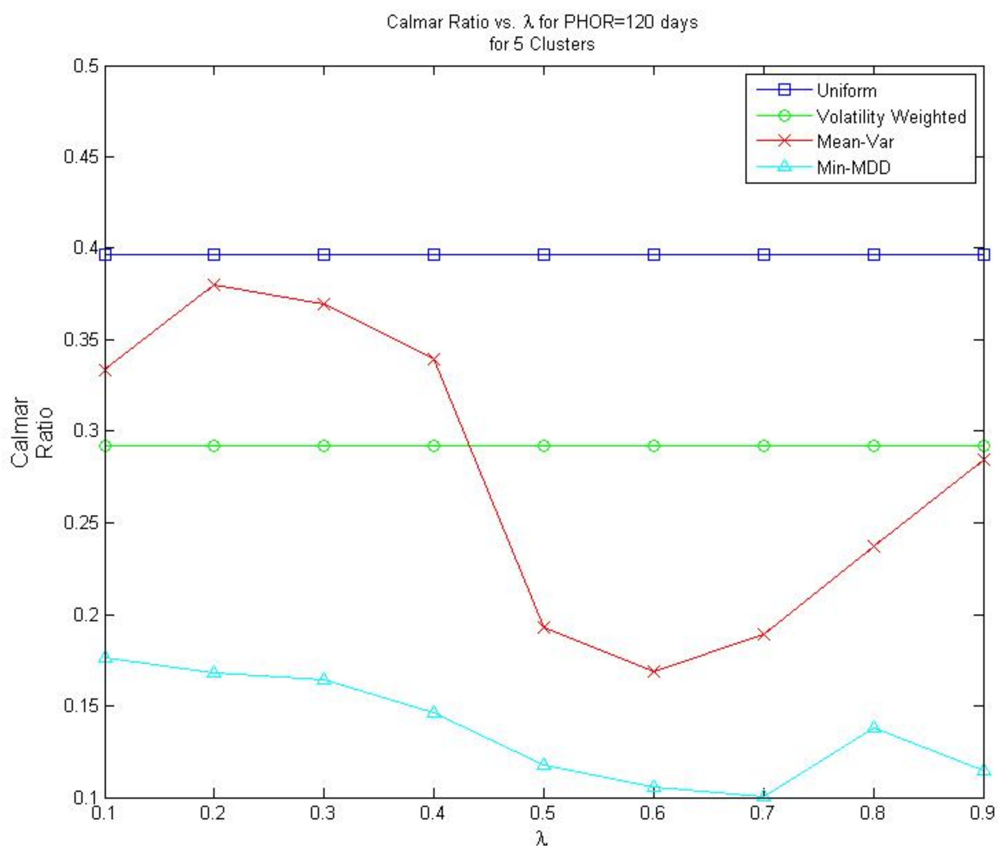

Sharpe Ratio vs. λ for PHOR=60 days
for 5 Clusters
Sharpe Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Calmar Ratio vs. λ for PHOR=60 days
for 5 Clusters
Calmar Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

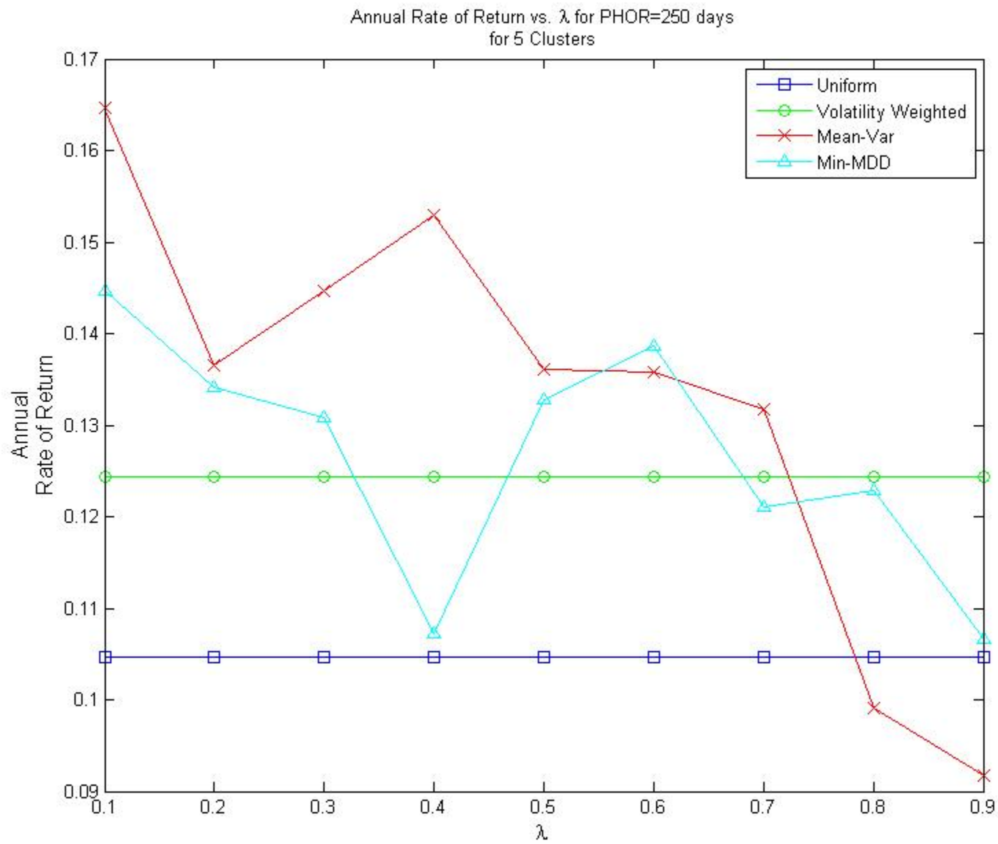
B.10 PHOR=120 days, 5-Clustering

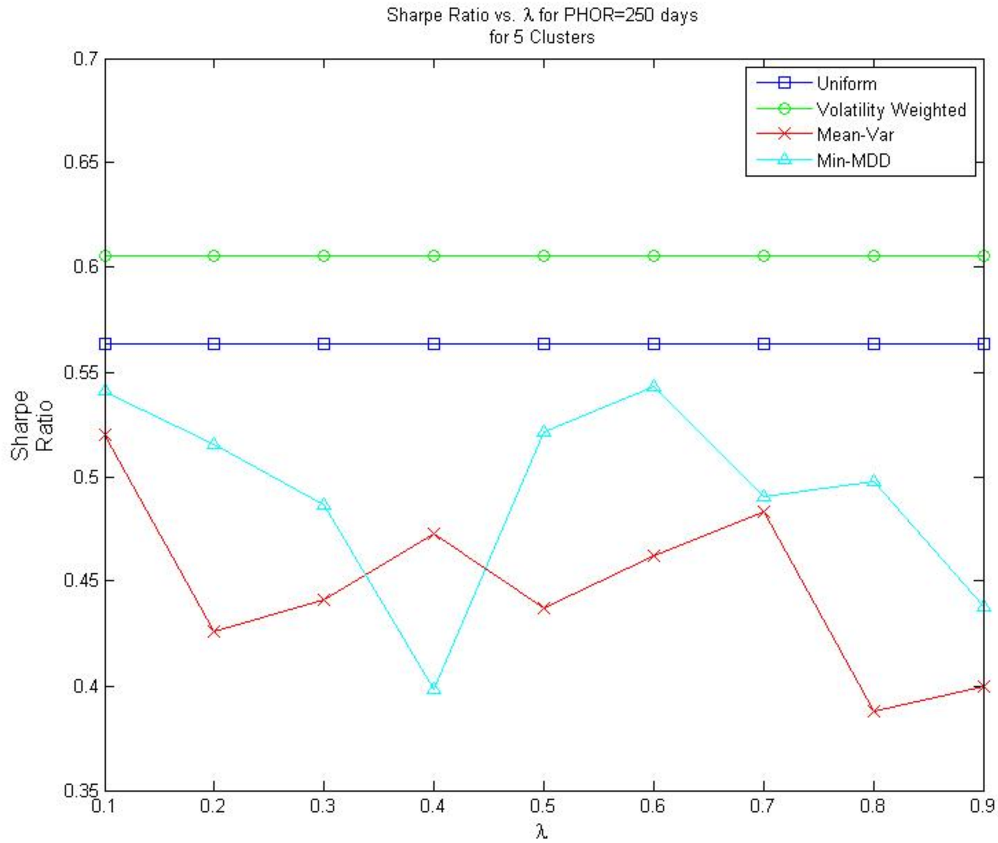


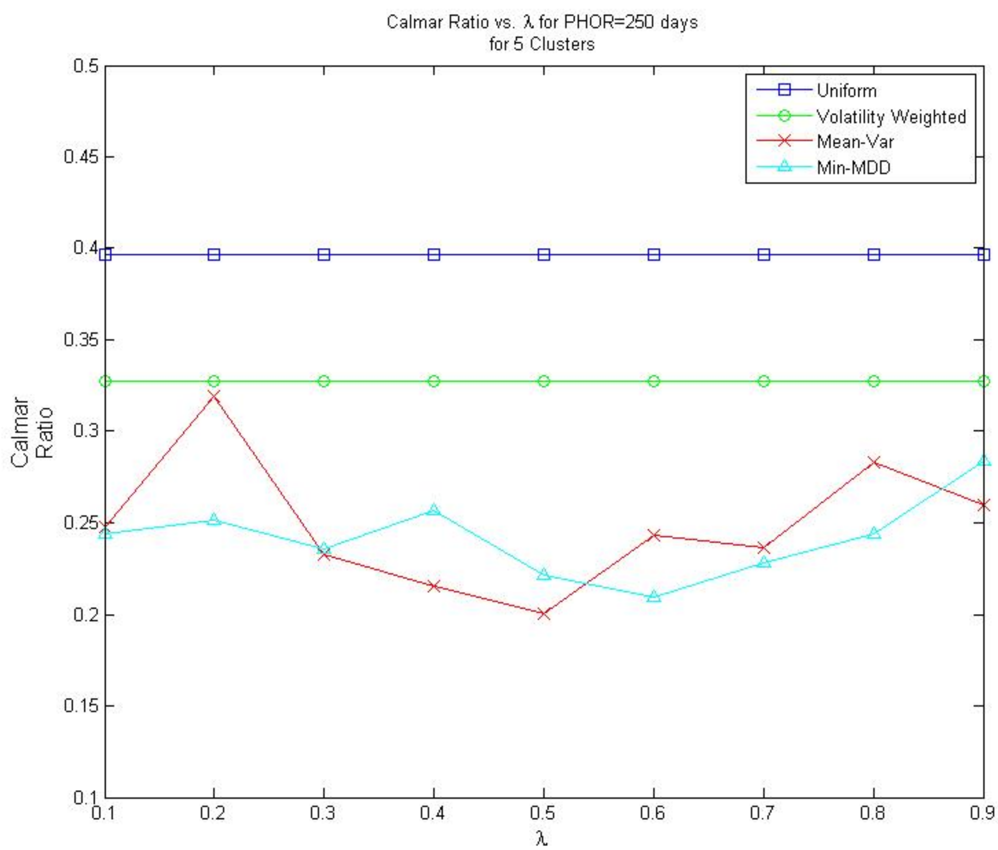

Sharpe Ratio vs. λ for PHOR=120 days
for 5 Clusters
Sharpe Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Calmar Ratio vs. λ for PHOR=120 days
for 5 Clusters
Calmar Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

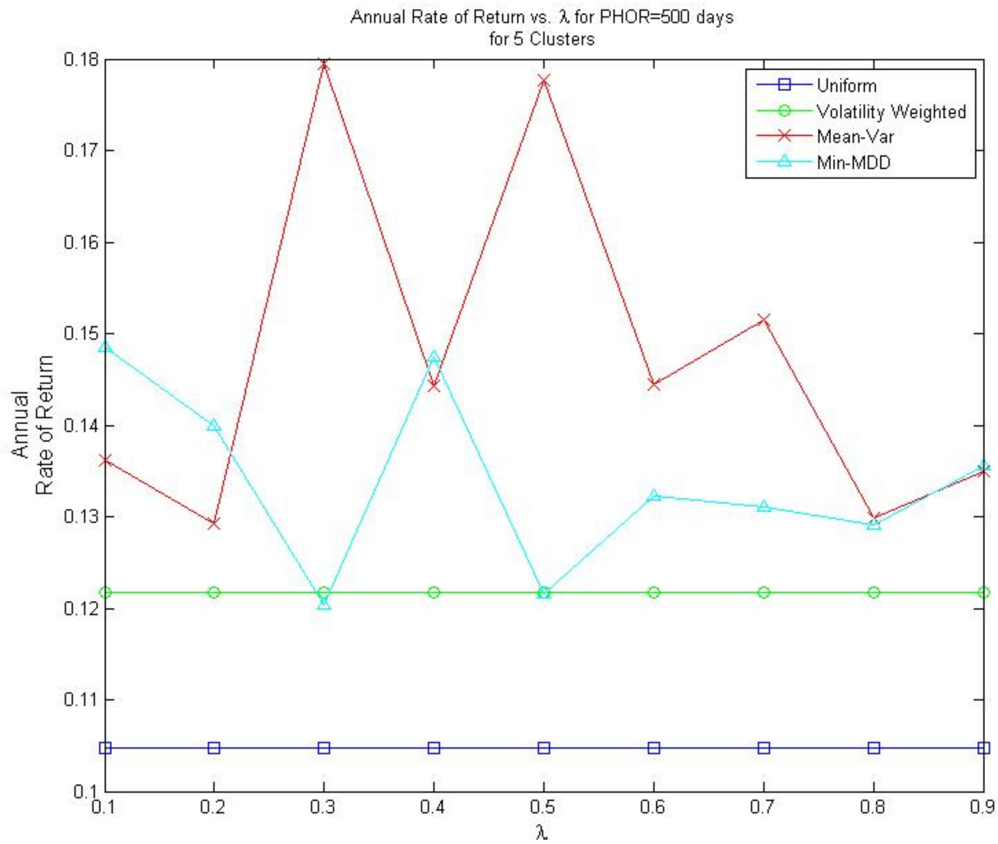
B.11 PHOR=250 days, 5-Clustering

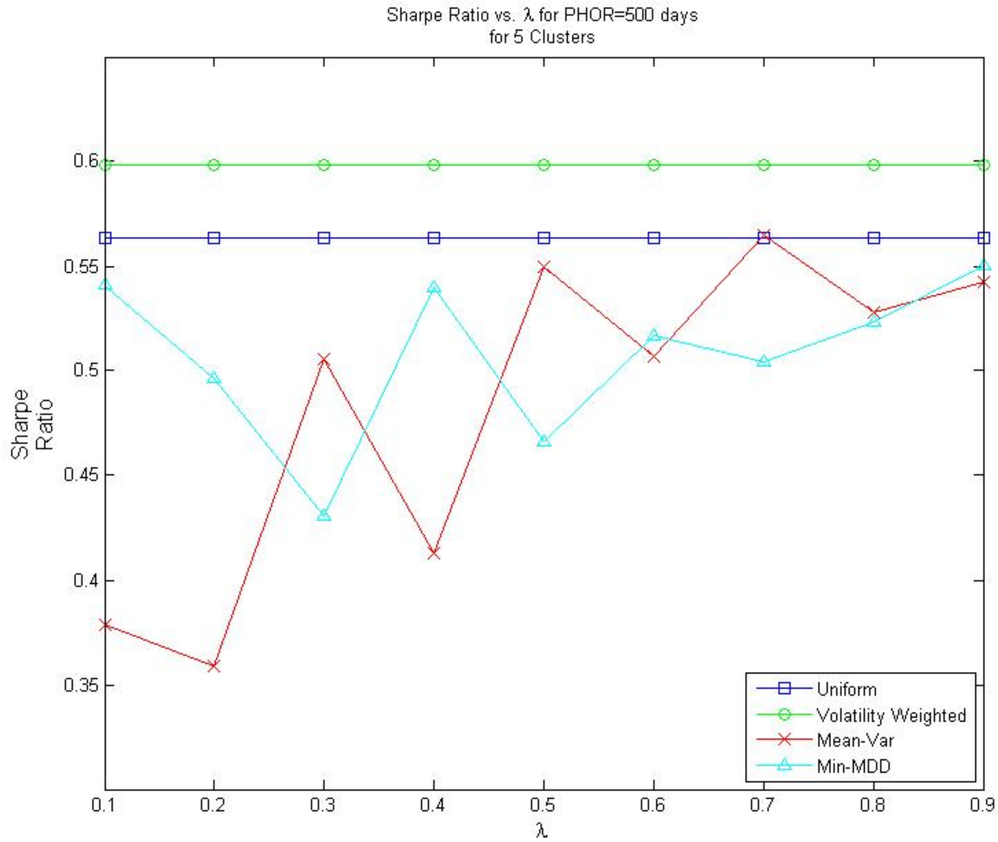


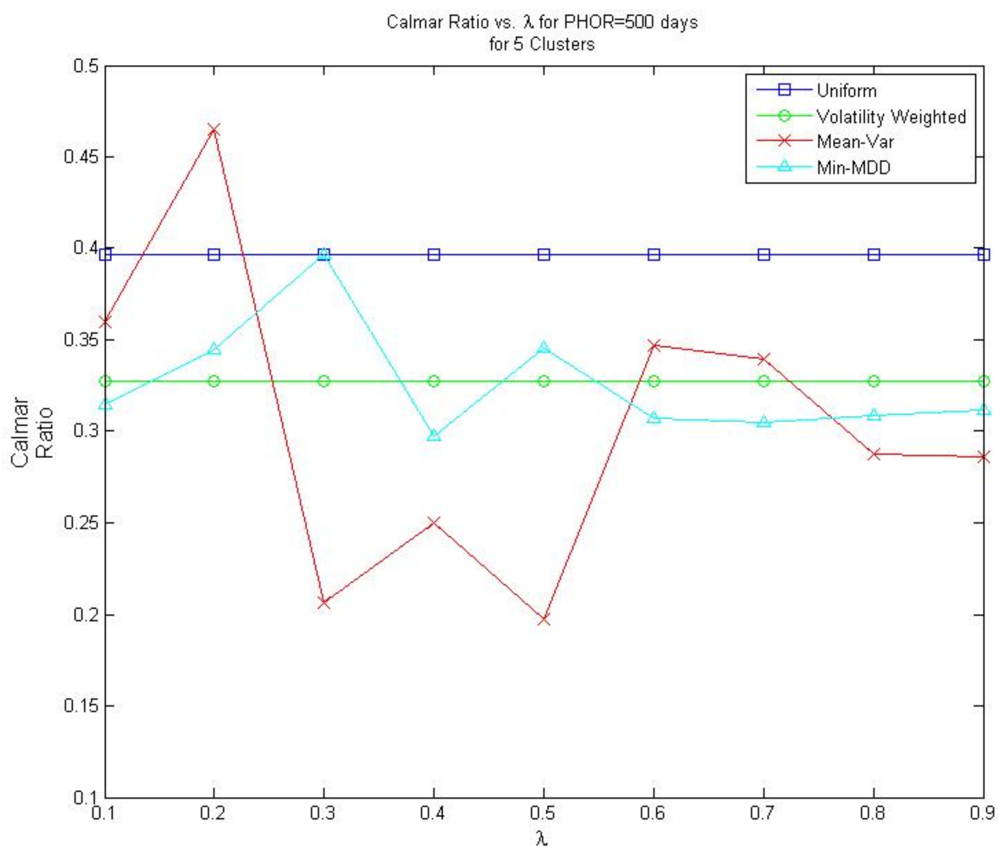

Sharpe Ratio vs. λ for PHOR=250 days
for 5 Clusters
Uniform
Volatility Weighted
Mean-Var
Min-MDD
Sharpe
Ratio
λ
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9


Calmar Ratio vs. λ for PHOR=250 days
for 5 Clusters
Uniform
Volatility Weighted
Mean-Var
Min-MDD
Calmar
Ratio
λ
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

B.12 PHOR=500 days, 5-Clustering




Sharpe Ratio vs. λ for PHOR=500 days
for 5 Clusters
Sharpe Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD


Calmar Ratio vs. λ for PHOR=500 days
for 5 Clusters
Calmar Ratio
λ
Uniform
Volatility Weighted
Mean-Var
Min-MDD