Summer 2021 COMP SCI 799 Research
University of Wisconsin-Madison

A Retail Investor's Tool for Portfolio Optimization and Risk Management

Jacob Lorenz
https://github.com/TheExplorativeBadger/portfolio-optimization

August, 2021

Abstract

Securities with an impressive overall historical track record may seem like intuitive investments, and investors may think they're diversifying by simply buying a few of the top performing securities instead of putting all their money into one. With large rises; however, comes the potential for equally as large of falls, and ignoring strategies for managing risk exposure has a non-insignificant probability of resulting in the investor losing a significant portion of their portfolio's value. Value at Risk is a risk measure used to understand the maximum expected loss with a given confidence interval, and Conditional Value at Risk, sometimes called Expected Shortfall, measures the expected loss in the worst case scenarios that exceed the Value at Risk threshold for the given confidence level. A valid risk management strategy is choosing portfolios based on minimizing the Value at Risk; however, minimizing Value at Risk can prove to be difficult mathematically in the many cases where it is not convex. Minimizing Conditional Value at Risk; however, is closely tied to reducing Value at Risk, and Conditional Value at Risk reduces nicely to a convex, stochastic optimization problem capable of being run efficiently over large data sets. Conditional Value at Risk is largely regarded as a more consistent measure of risk than Value at Risk, and the optimization process can be generalized outside of finance to problems involving the optimization of percentiles. The optimization of Conditional Value at Risk can be implemented programmatically with the help of other procedures for random scenario sampling, such as a Monte Carlo Simulation, and the optimization algorithm returns both the optimal portfolio allocations, as well as the portfolio's Value at Risk for a specified confidence level, making it an efficient and useful algorithm for the optimization tool to have in its set of capabilities.

Contents

1. Introduction

Investing in securities can be a great way for an individual to put their hard earned money to work and build a comfortable financial future. After taking the step to learn about the market and getting comfortable investing, the next question becomes: "Out of the thousands of possibilities, which securities should I invest in, and how much?". Beyond the strategy of picking securities at random or based on common household names, there exist many different quantitative analysis methods and metrics for portfolio optimization and risk management, and in the corporate world some of these methods constitute regulation that companies must abide by when making financial decisions on behalf of their clients and shareholders. Dating back to 1952, Economist Harry Markowitz developed a quadratic model for portfolio optimization based on the trade off between expected return and covariance risk between securities, which we will call the Mean-Variance (MV) approach. For this, Markowitz would go on to receive, among other awards, the Nobel Prize in 1990, and his work has been the foundation of evolving portfolio optimization strategies since. One shortcoming of the Mean-Variance approach is it inherently assumes past performance is a reliable indicator of future results, and it does not adequately account for the future randomness in financial markets. A security may have an impressive track record with a high overall average return; however, bad outcomes do happen and ignoring the worst-case performance scenarios can lead to devastating losses.

For financial companies, regulation imposes risk-management requirements in terms of percentiles of expected loss distributions. A common metric for this is the portfolio Value at Risk (VaR). VaR is determined in the context of a given confidence probability. The 90%-VaR, for example, is the loss where 10% of scenarios have losses exceeding the 90%-VaR. VaR has become a popular percentile metric due to its simplicity and intuitiveness; however, it has the potential for some mathematical shortcomings that are described in detail in (Krokhmal et al., 2001)[6], (Rockafellar, 1970)[10], and (Artzner et al., 1997, 1999)[1][2]. Namely, VaR is non sub-additive and in the case of discrete distributions, which is the case for the approach in this paper, VaR is non-convex. For these reasons, it is not conducive for our purposes to optimize over VaR; however, closely related to VaR, another common percentile measure of risk is called Conditional Value at Risk (CVaR). CVaR gives insight into the expected loss in the worst case scenarios, namely, CVaR is the conditional expectation of losses at least VaR. If the 95%-VaR of a distribution is the highest loss we can expect to lose with 95% confidence, 95%-CVaR is the expected loss in the 5% of scenarios not covered in the confidence interval. While VaR may not be suitable for our approach, Pflug (2000)[9] proved that CVaR is a coherant risk measure, and has the following properties: transition-equivariant, positively homogeneous, convex, and a few others beyond the scope needed for this paper. Minimizing the CVaR is closely related to minimizing VaR, and (Rockafellar and Uryasev, 2000)[11] proposes an approach for CVaR minimization, resulting in both the optimal portfolio and optimal portfolio's VaR, which will be followed in this paper.

The primary outcome of this research comes in the form of a programmatic tool that performs portfolio optimization according to the methodologies outlined in this paper. The tool uses the MV approach as a baseline, as well as the approach of optimizing over CVaR to generate respective optimal portfolios and draw comparisons. This paper acts as a supplemental description of the background, related work, mathematical formulation, methodology, and results of running the tool across all securities in the S&P500 and 14 of the top cryptocurrencies by market cap. We begin by discussing similar work in the field and the inspiration for this research, followed by a discussion of the data and data source used. We then formulate the optimization methods mathematically, describe the approach for simulation, and conclude with results and a discussion on the takeaways and future work.

2. Related/Similar Work

The core ideas, concepts, and methodologies used in this paper are not novel by any means; however, the results and implementation are. (Rockafellar and Uryasev, 2000)[11] first proposed the Conditional Value at Risk metric and the outline of an optimization algorithm, and (Krokhmal et. al, 2001)[6] and Rockafellar and Uryasev (2001)[12] expanded upon it using the same ideas and a few new features. All of these papers rely on additional mathematical conclusions cited from Rockafellar (1970)[10], Shor (1985)[13], Uryasev (1995)[14], Kan and Kibzun (1996)[4], Kast et. al (1998)[5], and Mauser and Rosen (1999)[7], and the proofs / derivations for the theorems and conclusions used in this paper can be found in those original works. The approach and methodologies used in this paper are drawn from the ideas in Rockafellar and Uryasev (2000)[11], Krokhmal et. al (2001)[6], Rockafellar and Uryasev (2001)[12], and Optimization Methods in Finance (2nd) by Cornuéjols, Peña, and Tütüncü[3].

3. Dataset

The data used in this research is procured from Alphavantage using their Daily Adjusted Time Series (https://www.alphavantage.co/documentation/dailyadj) and Daily Digital Currency (https://www.alphavantage.co/documentation/currency-daily) API endpoints. A list of security ticker symbols for the securities in the S&P500 and top 14 Cryptocurrencies by market cap was gathered, and the tickers are used as keys to access the data of the securities via the APIs. The data comes in the form of daily time series points gathered over the last 20 years, or lifetime if the security history is younger than 20 years old. Daily data points and return were considered on the basis of closing price between successive trading sessions.

4. Model

4.1 Mean-Variance

In general, the idea of the Mean-Variance approach proposed by Markowitz is to use a trade off between the expected return versus the expected risk (variance) by using a regularization parameter to give significance to the risk versus reward weight.

Let $\mathbf{x} = (x_1, x_2, \dots, x_n)$ be a portfolio of n securities with x_i representing the fraction of the portfolio allocated to security i. Then, we have

$$\sum_{i=1}^{n} x_i = 1, \quad x_i \geq 0 \ \ \forall i \in (1, n)$$

To promote diversification, the tool allows for an arbitrary maximum individual security allocation limit, δ, to be set to any value the investor desires.

$$x_i \leq \delta \ \ \forall i \in (1, n)$$

Let $\mathbf{Cov}_{i,j}$ be the covariance of the expected return between securities i and j, defined by

$$\mathbf{Cov}_{i,j} = \frac{1}{d} * \sum_{k=1}^{d} (r_{i,k} - \mathbf{Ret}_i) * (r_{j,k} - \mathbf{Ret}_j) = \mathbf{Cov}_{j,i}$$

where d = number of daily return data points, $r_{i,k}$ = daily return of security i on day number k calculated between successive close prices, and $\mathbf{Ret}_i$ = average daily return of security i. Then, given a regularization parameter λ, the optimal portfolio can be calculated using quadratic optimization for the objective function

$$\min_{x_i} \quad \lambda * \sum_{i=1}^{n} \sum_{j=1}^{n} (x_i * \mathbf{Cov}_{i,j} * x_j) - \sum_{i=1}^{n} (\mathbf{Ret}_i * x_i)$$

After obtaining the optimal portfolio allocation vector, $\mathbf{x}^*$, expected portfolio daily return can be calculated with $\mathbf{x}^* * \mathbf{Ret}$.

4.2 Conditional Value at Risk

As mentioned in the summary of related work, the model used for CVaR optimization in this paper is derived primarily from Rockafellar and Uryasev (2000)[11], Krokhmal et. al (2001)[6], Rockafellar and Uryasev (2001)[12], and Optimization Methods in Finance (2nd) by Cornuéjols, Peña, and Tütüncü[3]. For a complete walkthrough of the method's derivation, the reader is encouraged to examine those works independently. Here we will present the core theorems and derivations necessary to adequately justify the approach.

Let $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ be a portfolio of n securities with x_i representing the fraction of the portfolio allocated to security i. Then, we have

$$\sum_{i=1}^{n} x_i = 1, \qquad x_i \geq 0 \ \ \forall i \in (1, n)$$

Let $f(x, y)$ be the loss of the decision vector x, and the random vector y_n. For each potential portfolio, x, the loss $f(x, y)$ is a random variable with distribution induced by the distribution of y. Assume that the underlying distribution has density, denoted $p(y)$. $p(y)$ would naturally be 0 for all y if y is continuous; however, for the sake of calculation we approximate the probability density with a Monte Carlo simulation to sample the distribution (JP Morgan, 1996)[8]. Then, the cumulative distribution function of $f(x, y)$, $\Psi(x, \alpha)$, gives the probability of $f(x, y)$ not exceeding α, and is given by

$$\Psi(x, \alpha) = \int_{f(x,y)\leq\alpha} p(y)dy$$

We assume, for simplicity, that $\Psi(x, \alpha)$ is everywhere continuous with respect to α. Then, the β-VaR, $\alpha_\beta(x)$, and β-CVaR, $\phi_\beta(x)$, values for $f(x, y)$ and any probability level β in (0,1) are given by

$$\alpha_\beta(x) = min\{\alpha \in \mathbb{R} : \Psi(x, \alpha) \geq \beta\}$$

and

$$\phi_\beta(x) = \frac{1}{1-\beta} \int_{f(x,y)\geq\alpha_\beta(x} f(x, y)p(y)dy$$

$\alpha_\beta(x)$ comes out as the left endpoint of the interval consisting of all α such that $\Psi(x, \alpha) = \beta$, and $\phi_\beta(x)$ comes out as the expectation of the loss of x conditional to that loss being $\alpha_\beta(x)$ or greater. Rockafellar and Ursayev (2000)[11] present a key characterization of $\alpha_\beta(x)$ and $\phi_\beta(x)$ in terms of a function F_β defined as

$$F_\beta(x, \alpha) = \alpha + \frac{1}{1-\beta} \int_{y\in\mathbb{R}^m} [f(x, y) - \alpha]^+ p(y)dy$$

where

$$[t]^+ = \begin{cases} t, & \text{when } t > 0 \\ 0, & \text{when } t \leq 0 \end{cases}$$

They also present two theorems that guide the optimization derivation and hold a discussion on some of the theoretical implications and assumptions of the theorems. For brevity's sake, we will omit a discussion on the consequences, and just provide the theorems and outcome.

Theorem 1: *As a function of α, $F_\beta(x, \alpha)$ is convex and continuously differentiable. The β-CVaR of the loss associated with any $x \in X$ can be determined from the formula*

$$\phi_\beta(x) = \min_{\alpha\in\mathbb{R}} F_\beta(x, \alpha)$$

In this formula, the set consisting of the values of α for which the minimum is attained, namely

$$A_\beta(x) = arg \min_{\alpha \in \mathbb{R}} F_\beta(x, \alpha),$$

is a nonempty closed bounded interval, and the β-VaR of the loss is given by

$$\alpha_\beta(x) = \textit{left endpoint of } A_\beta(x)$$

In particular, on always has

$$\alpha_\beta(x) \in arg \min_{\alpha \in \mathbb{R}} F_\beta(x, \alpha) \quad \textit{and} \quad \phi_\beta(x) = F_\beta(x, \alpha_\beta(x))$$

To approximate the integral in the original equation of $F_\beta(x, \alpha)$, we can sample the probability distribution using the Monte Carlo simulation and generate a collection of vectors $y_1, \dots, y_q$, then the approximation is given by

$$\tilde{F}_\beta(x, \alpha) = \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} [f(x, y_k) - \alpha]^+$$

which is convex and piecewise linear with respect to α. Although not differentiable, it can be easily minimized with basic linear programming. Furthermore,

Theorem 2: *Minimizing the β-CVaR of the loss associated with x over all $x \in X$ is equivalent to minimizing $F_\beta(x, \alpha)$ over all $(x, \alpha) \in X \times \mathbb{R}$, in the sense that*

$$\min_{x \in X} \phi_\beta(x) = \min_{(x,\alpha) \in X \times \mathbb{R}} F_\beta(x, \alpha)$$

where, moreover, a pair $(x^*, \alpha^*$ achieves the second minimum if and only if x^* achieves the first minimum and $\alpha^* \in A_\beta(x^*)$. In particular, therefore, in circumstances where the interval $A_\beta(x^*)$ reduces to a single point, the minimization of $F(x, \alpha)$ over $(x, \alpha) \in X \times \mathbb{R}$ produces a pair (x^*, α^*), not necessarily unique, such that x^* minimizes the β-CVaR and α^* gives the corresponding β-VaR. Furthermore, $F_\beta(x, \alpha)$ is convex with respect to (x, α) and $\phi_\beta(x)$ is convex with respect to x, when $f(x, y)$ is convex with respect to x, in which case, if the constraints are such that X is a convex set, the joint minimization is an instance of convex programming.

The key takeaway from Theorem 2 is that the optimization problem now falls under the category of stochastic optimization due to the presence of the expectation in the definition of $F_\beta(x, \alpha)$. Rather than operate on the β-CVaR directly, which can be tricky due to its mathematical properties, we can use the convex function $F_\beta(x, \alpha)$. Theorem 2 allows us to use stochastic programming techniques for the minimization of β-CVaR.

Let $y = (y_1, \dots, y_n)$ be the random vector of returns for each security in a given simulation iteration. Then, the loss of the portfolio over a single iteration, defined as a percentage, is given by

$$f(x, y) = -[x_1 y_1 + \cdots + x_n y_n] = -x^T y$$

The objective function of the optimization problem is connected to β-VaR and β-CVaR, and is given by

$$F_\beta(x, \alpha) = \alpha + \frac{1}{1-\beta} \int_{y \in \mathbb{R}} [-x^T y - \alpha]^+ p(y) dy$$

Which, approximated linearly, becomes

$$\tilde{F}_\beta(x, \alpha) = \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} [-x^T y_k - \alpha]^+$$

By using auxiliary variables, u_k for $k = 1, \dots, q$, the approximation can be reduced to convex programming, and is equivalent to minimizing

$$\alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} u_k$$

subject to

$$u_k \geq 0 \quad \text{and} \quad x^T y_k + \alpha + u_k \geq 0 \qquad \text{for } k = 1, \dots, q$$

Finally, we add a constraint imposing a requirement that only portfolios with an expected return of at least a given amount, R, will be considered. Let $\mu(x)$ denote the mean loss associated with portfolio x. Let m be the mean of the loss distribution y. Then

$$\mu(x) = -x^T m$$

and our constraint is

$$\mu(x) \leq -R$$

In conclusion of the approach for optimization over β-CVaR, we have the following convex, stochastic, optimization model:

$$\min_{x,\alpha} \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} u_k$$

subject to

$$\mu(x) \leq -R, \quad x^T y_k + \alpha + u_k \geq 0, \quad x_i \geq 0, \quad u_k \geq 0 \quad \forall \;\; i \in 1, \ldots, n, \quad k \in 1, \ldots, q$$

The model returns the optimal portfolio, x^*, abiding by the constraints and the optimal portfolio's β-VaR, α^*. The portfolio's expected percentage return can be found with $x^{*T}m$ where m is the expected return vector of all securities based on the Monte Carlo simulation results.

5. Approach

5.1 Mean-Variance

For the Mean-Variance optimization, we begin by obtaining the daily time series data for up to the last 20 years of history for the securities in S&P500 and top 14 cryptocurrencies by market cap. We calculate the daily returns as a percent change between successive closing prices for all trading sessions across all securities, and found the average, variance, and standard deviation for each security. We next calculate the covariance matrix for the daily returns of all securities. Each security is also tagged with an industry category. In addition to calculating the average return and covariance between individual securities, we sort the securities by industry, aggregate all the data, and calculate the average industry daily return and industry covariance matrix for use in the optimization process.

After aggregating all the data and performing the initial calculations, we move on to the optimization process. We take a 2-step approach for the sake of promoting diversification; this could easily be extended into alternative cycles and different strategies can be used to experiment. Step one is optimizing the quadratic model from Section 4.1 using the industry x industry data and a maximum industry allocation limit of 0.2. We perform this optimization 50 times with λ values between $10^{-}3$ and 10^4 and use the industry allocation distribution from the highest expected return trial to proceed. After obtaining the highest returning industry allocation, we perform a second round of the same quadratic optimization, but with the data of individual securities. For each industry chosen in the optimal industry allocation, we perform isolated optimization over all securities belonging to the respective industry, again with a .2 allocation limit. When combined with the industry allocation limit of .2, this effectively makes the maximum allocation of any given security .04, or, 4% of the total portfolio. We do this 50 times across the same λ range, and take our final optimal portfolio to be that with the highest expected return. This is another matter of taste, the tool can easily be tweaked to return the portfolio with an other measure's extrema instead of highest return.

5.2 Conditional Value at Risk

For the optimization over CVaR, we take a slightly different approach using a Monte Carlo Simulation. Using the historical daily time series data for each individual security, we find the periodic daily return for each day, q, given by

$$R_{i,q} = ln(\frac{C_{i,q}}{C_{i,q-1}})$$

Where $C_{i,q}$ = close price of security i on day q. From this, we calculate the average daily periodic return, PAvg_i and the periodic variance, PVar_i for each security and use them to calculate the security's drift, defined as

$$D_i = \text{PAvg}_i - \frac{\text{PVar}_i}{2}$$

Using the calculated drift and standard deviation, and price per unit of each security, we begin the Monte Carlo simulation. We run the simulation 1,000 times for each security. For each round, we simulate 90 days in the market using the following 2-step formula to simulate the price change of a single day:

1. Obtain a random seed value based on the periodic standard deviation and norm inverse of the quantile of a random number between 0 and 1 on a standard normal distribution

$$RV = \sigma_i * \Phi^{-1}(\text{rand}(0, 1))$$

2. Simulate the current day's (s) price movement

$$P_{s+1} = P_s * e^{(D_i + RV)}$$

After performing 1,000 simulations across all securities, each for a simulated period of 90 days, we obtain the 90-day overall loss. It's important to note that the loss is simply the opposite of gain. If a simulation has a loss of -.1, that is actually saying the simulation resulted in a gain of 10%. A loss histogram with 1000 even-width buckets is created for each security and we count the number of simulations with a 90-day loss within the bounds of each bucket to populate the counts. When you consider the total number of simulations and occurrence count for each bucket in the loss histogram, you can approximate the probability of each simulated outcome discretely.

With the histograms created and expected loss of each security across all scenarios calculated, we move on to the optimization process. Once again, this process uses a 2-step approach, and is one of many different ideas that could be followed. We use the convex optimization model outlined in

Section 4.2 to determine the optimal portfolio at the 95% VaR and CVaR threshold, with the restriction that the expected loss of the overall portfolio is smaller than 0, i.e. we only take portfolios into consideration that are expected to have positive return based on the simulation results. After obtaining the allocation vector of the optimal portfolio, we identify the securities which have an allocation percentage of at least 1%. Any security that has an allocation percentage less than 1% is removed from consideration, and the optimization is run one more time over all securities with an original allocation of at least 1%. Once again, the resulting allocation vector is analyzed for the allocations of at least 1%, and these resulting securities are included among the final optimal portfolio at the allocation percentage specified. Any floating point portfolio percentage left over after pruning the final results can be thought of as holding a cash position rather than spending every last dollar on small, left-over positions in more securities.

6. Results

6.1 Mean-Variance

After following the steps outlined for Mean-Variance Optimization in Section 4.1 and 5.1, the tool produces a decision vector, x, which contains the percentage allocation of each chosen security in the optimal portfolio. A breakdown of this allocation distribution can be found in Table 1 of Appendix A. The portfolio contains 27 securities and has a past performance average daily return of .3113%, extrapolating to an annual return of 211% and 90 day return of 32.28%. The 95%-VaR and 95%-CVaR values were found for the optimal portfolio to be 2.13% and 5.41%, respectively, for a 90 day simulation.

For a further breakdown of analysis on the individual securities' time series, historical return, covariance, and other optimization metrics, see Appendix B.

6.2 Conditional Value at Risk

After following the steps outlined for Conditional Value at Risk Optimization in Section 4.2 and 5.2, the tool produces a decision vector, x, which contains the percentage allocation of each chosen security in the optimal portfolio. A breakdown of this allocation distribution can be found in Table 2 of Appendix A. The portfolio contains 33 securities and has an expected 3 month return of 15.54% with a 95%-VaR and 95%-CVaR of -.193% loss, which translates to a .193% gain.

For a further breakdown of analysis on the metrics involved with CVaR Optimization, see Appendix C.

6.3 Discussion

Admittedly, it was a bit surprising when the 95%-VaR and 95%-CVaR values came back as the same value for CVaR Optimization; however, there could be some explanations for it. Due to processing time and the discrete nature of the study, the loss distribution histograms were limited to 1000 buckets capturing the entire spread from minimum to maximum loss for all securities. For this reason, the bucket widths captured a range of 2.4% each, and each occurrence of the bucket was generalized to the mean value of the bucket. Perhaps the worst performing cases all fell into the upper end of the histogram and the 95 percentile values fell into the same loss bucket as the 100 percentile values. Further compounding on this, is the constraint that limits the portfolios considered to those with an expected loss ≤ 0. Perhaps this constraint limits the feasible set to candidate portfolios where the worst-performing cases are limited and bundled together. The result of the optimal CVaR portfolio having a 95%-VaR of a .193% gain gives confidence in the model, given that the constraint was set to only consider portfolios with an expected return ≥ 0. The extrapolated annual return based on the simulated 3 month expected return of 15.54% comes to 78.21%. This is clearly less than the MV expected annual return of 211%; however, the MV portfolio comes with a higher degree of risk and the 95%-VaR and 95%-CVaR are both higher losses than the CVaR optimized portfolio, which makes sense. One might read these results and say they're willing to take the chance of 5% loss for higher gains; however, the investor would also be cautioned that these metrics are based primarily on historical data and certainly do not capture the true scope of random fluctuation. They are models to help understand an idea of potential outcomes, and faith ought to be placed with the models that best account for uncertainty, in which case the portfolio optimized over CVaR should be regarded as being more considerate of the random element. In any case, more work and refinement would certainly not hurt, and continued refactoring of the tool only makes it more efficient and capable.

7. Conclusions and Future Work

The methodologies outlined in this paper have guided the reader on the theoretical and functional steps taken to build the underlying portfolio optimization tool. The data and results presented is on the basis of the securities in the S&P500 and top 14 cryptocurrencies by market cap; however, additional securities can easily be added to the collection. The tool is capable of handling variable maximum and minimum allocation limits, and it is built such that optimal portfolios can be returned for different metric extrema in the case of the Mean-Variance approach. The tool is capable of running a variable number Monte Carlo simulations for variable length simulation duration, and the histograms describing the probability of various outcomes can be tuned to have variable bucket widths. Everything about the tool was made with customization and experimentation in mind, up to that allowed within the model formulations.

Conditional Value at Risk, or the convex function the derivation reduced it to, proves to be a valuable metric to optimize over. As an anecdotal observation, securities with high return potential also have some of the highest loss potential. Looking at astronomical historical gains can blind the enthusiastic investor to the risks of significant portfolio loss in a non-insignificant percentage of outcomes. Obtaining a collection of well-performing securities, then minimizing risk exposure with optimization over CVaR, while providing a desired return minimum, can be a step in the right direction. Strictly speaking from the perspective of finance, finding the best performing investments while reducing your risk exposure and expected loss in the worst case scenarios is indeed financial optimization, and CVaR optimization does just that. One disclaimer worth making again, is past performance is not a predictor of future results. Compared to MV optimization, which uses only the past performance, CVaR optimization takes a step in the right direction by using the past data to generate a sequence of future outcomes based on random variables used to mimic the randomness of the market; however, even this is still based on past performance, and the mimicked randomness surely does not fully capture the true spectrum of causes of market fluctuation. When it comes down to it, randomness is just that, randomness, and the best attempts at finding a pattern to justify the randomness usually result in probability distributions and expectations. So, with the premise that we cannot truly understand everything involved in randomness, the next best solution is to maximize the rewarding probability and develop measures around limiting the exposure probability, and in that sense CVaR is a very viable candidate metric.

The tool and its capabilities continues to be a work in progress. It currently has capabilities for Mean-Variance and CVaR optimization; however, there are surely more metrics and objective functions in the quantitative finance world that have not been explored, implemented, and used for comparison in the tool. Finding thoroughly developed optimization strategies and implementing them programmatically is no short task; however, one by one, the addition of further refined algorithms to the tool builds out its capabilities and the confidence in results. Another area to explore in the near future is how these results change when securities from other markets are considered. As mentioned, the tool is built such that the user is easily able to add in lists of security tickers from other markets, the only restricting factor is processing time; however, this again can be tuned by specifying variable simulation duration and number of iterations. Adding additional markets also allows for the introduction of a third round of MV optimization, by first optimizing over market allocation, then by industry, then security. Furthermore, running sample optimizations using a variety of parameter values would allow the user of the tool to understand how each parameter tweaks the recommendations, and they can then determine the appropriate set of parameters for their investment strategy. Thoroughly exploring the configurable parameters is ongoing work.

One of the more complex yet interesting ideas for future work on the tool is the idea of a programmatic feedback mechanism. The idea would be to run the optimization over multiple iterations, finding optimal portfolios, and using the optimal allocation of each iteration on a testing phase of real data to determine the true performance of that portfolio on a set time frame in the past. The results could then feedback via a set of parameters to tune weights associated with this hypothetical algorithm. A technical place to start would be finding existing applications of quantitative finance mixed with neural networks. In any case, the concept of including machine learning capabilities with the optimization tool is intriguing and would allow for more dynamic decision making.

In conclusion, a mathematically and technically adept reader would be able to use the models and approaches outlined in this paper, derived from literature on financial optimization, to build their own custom portfolio optimization tool. Or they could use the one hosted at http://www.github.com/TheExplorativeBadger/portfolio-optimization. Following the steps for data acquisition, transformation, aggregation, and simulation would allow them to end up at the point of being ready to optimize. Implementing the models via solvers such as Ipopt or Gurobi would allow them to determine an optimal portfolio allocation constrained by their individual parameters.

A. Optimal Portfolio Allocations

In Table 1 below, we give the resulting security allocation of the optimal portfolio determined by the Mean-Variance optimization approach.

Ticker	Name	Allocation
ALGN	Align Technology	3.996%
AMT	American Tower Corp.	3.999%
ANET	Arista Networks	3.992%
BKNG	Booking Holdings Inc	3.999%
CZR	Ceasars Entertainment	3.999%
CBRE	CBRE Group	3.999%
CMG	Chipotle Mexican Grill	.0273%
CCI	Crown Castle	3.941%
DEX	DexCom	3.999%
ENPH	Enphase Energy	3.999%
EQIX	Equinix	3.999%
ETSY	Etsy	3.999%
EXR	Extra Space Storage	.056%
HLT	Hilton Worldwide Holdings	3.969%
ISRG	Intuitive Surgical Inc	3.999%
LH	Laboratory Corp. of America Holding	3.999%
PAYC	Paycom	3.999%
PYPL	PayPal	3.998%
REGN	Regeneron Pharmaceuticals	3.999%
SBAC	SBA Communications	3.999%
NOW	ServiceNow	3.999%
TSLA	Tesla Inc.	3.999%
DOGE	Dogecoin	3.999%
THETA	Theta	3.999%
DOT	PolkaDot	3.999%
BNB	BNB	3.999%
LINK	Link	3.999%

Table 1: The security allocation of the optimal portfolio for MV optimization

In Table 2 below, we give the resulting security allocation of the optimal portfolio determined by the Conditional Value at Risk optimization approach.

Ticker	Name	Allocation
GOOG	Alphabet Inc. (Class C)	4.986%
ANET	Arista Networks	2.730%
AZO	AutoZone Inc.	3.237%
AVGO	Broadcom Inc.	3.833%
CARR	Carrier Global	4.999%
CTLT	Catalent	3.893%
CDW	CDW	4.664%
CHTR	Charter Communications	4.999%
CMG	Chipotle Mexican Grill	4.196%
DXCM	DexCom	1.459%
DPZ	Domino's Pizza	2.326%
FB	Facebook Inc.	1.191%
FRC	First Republic Bank	1.616%
FLT	FleetCor Technologies Inc	3.978%
FBHS	Fortune Brands Home and Security	3.623%
GNRC	Generac Holdings	3.484%
GWW	Grainger (W.W.) Inc.	1.215%
IR	Ingersoll Rand	1.819%
KEYS	Keysight Technologies	4.816%
NVR	NVR Inc.	2.569%
NXPI	NXP Semiconductors	2.328%
OTIS	Otis Worldwide	4.999%
PAYC	Paycom	4.999%
NOW	ServiceNow	3.710%
TDY	Teledyne Technologies	1.434%
TMO	Thermo Fisher Scientific	2.928%
BTC	Bitcoin	3.056%
ETH	Ethereum	1.404%
THETA	Theta	2.211%
ADA	Cardano	1.518%
DOT	PolkaDot	1.481%
LINK	Link	1.490%
XMR	XMR	1.651%

Table 2: The security allocation of the optimal portfolio for CVaR optimization

B. Analysis on Mean-Variance Optimization

In this Appendix, we show some metrics and analysis conducted during the Mean-Variance Optimization Process.

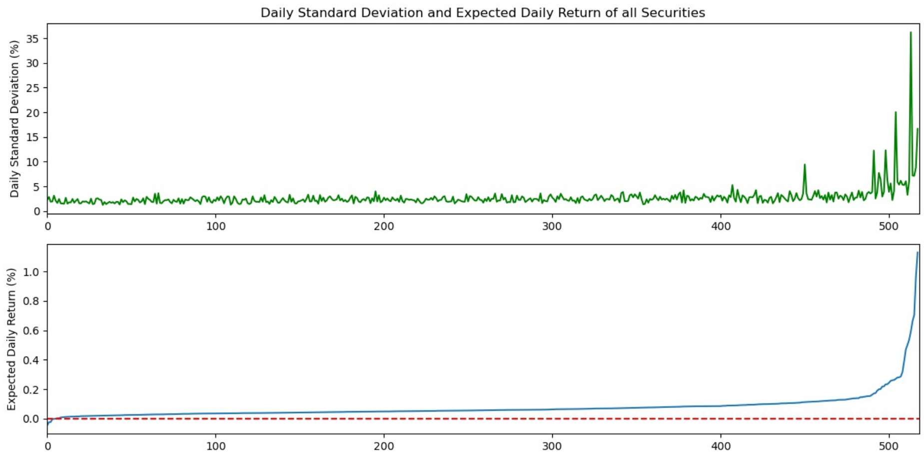

Figure 1: Daily Return % and Daily Return Standard Deviation % of all securities sorted by Daily Return %

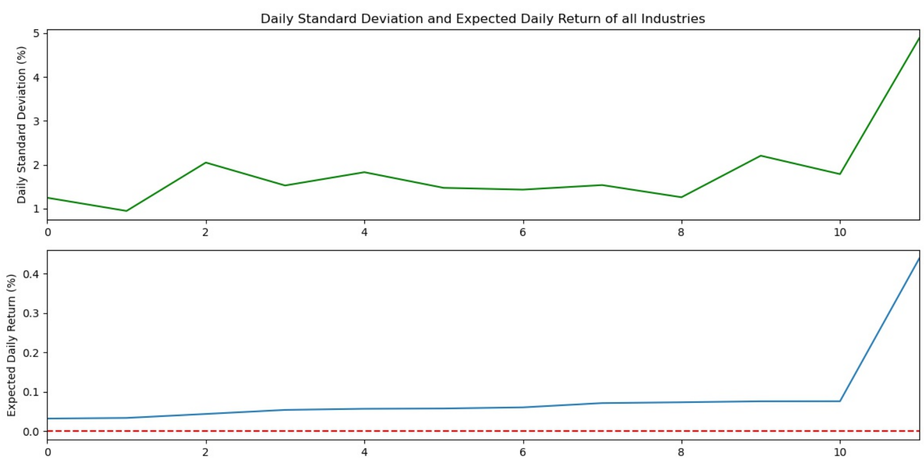

Figure 2: Daily Return % and Daily Return Standard Deviation % of all industries sorted by security Daily Return %

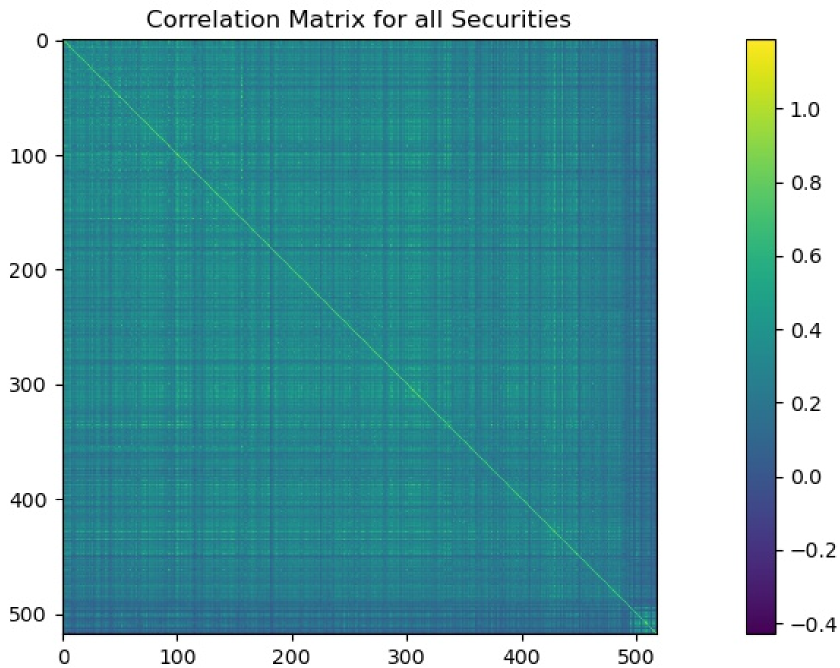

Figure 3: Correlation Matrix for security by security comparison

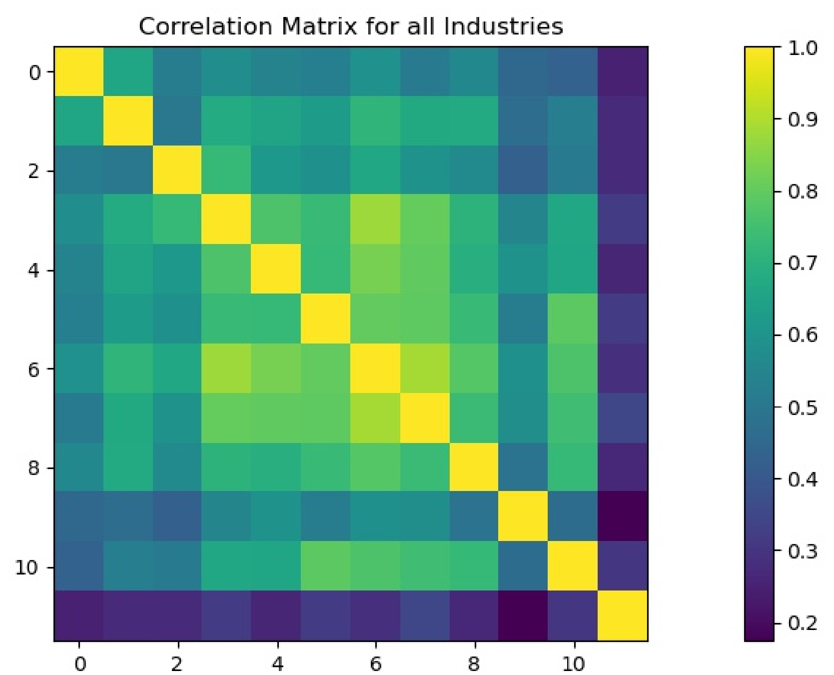

Figure 4: Correlation Matrix for industry by industry comparison

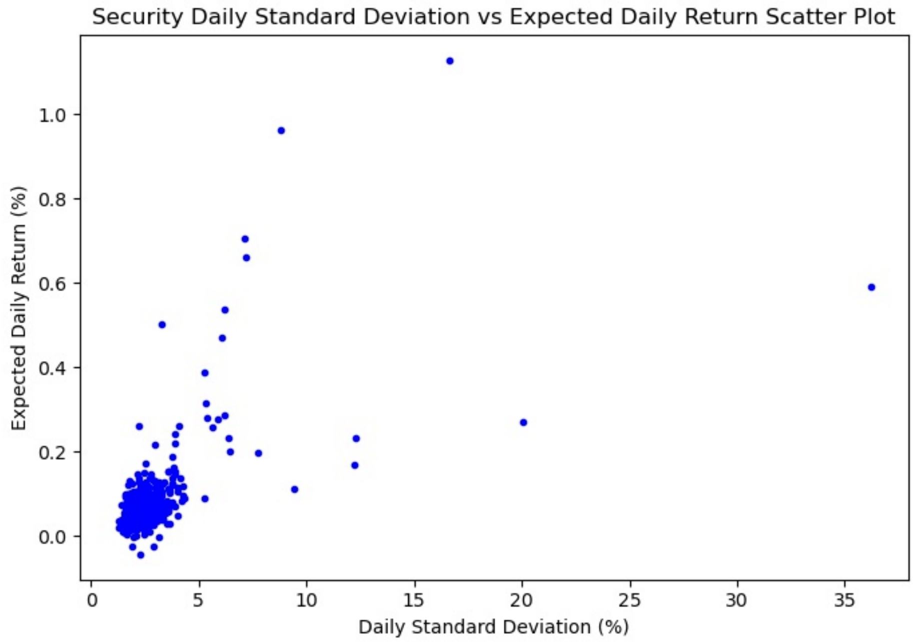

Figure 5: Individual security daily return and standard deviation scatter plot

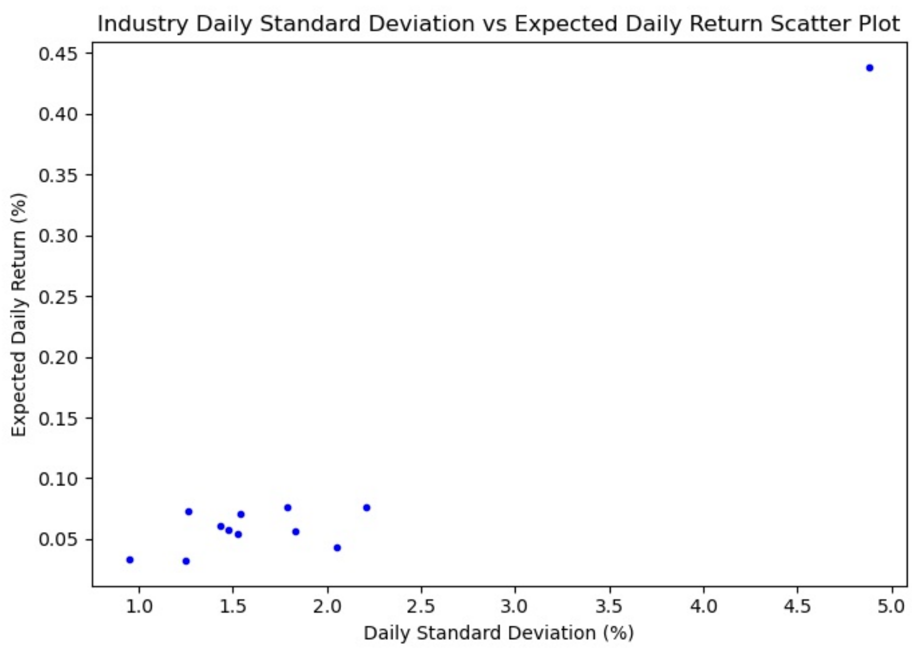

Figure 6: Industry daily return and standard deviation scatter plot

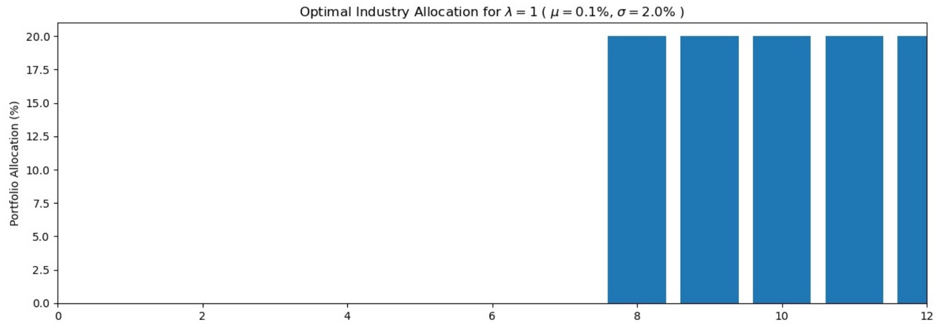

Figure 7: Optimal allocation percentages for industries with a regularization parameter of 1

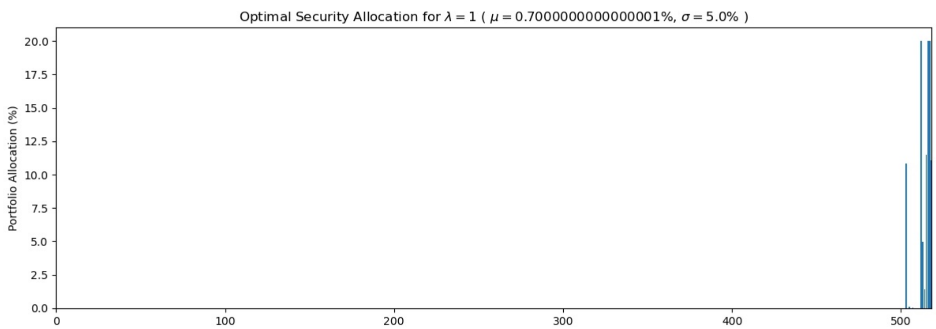

Figure 8: Optimal allocation percentages for securities with a regularization parameter of 1

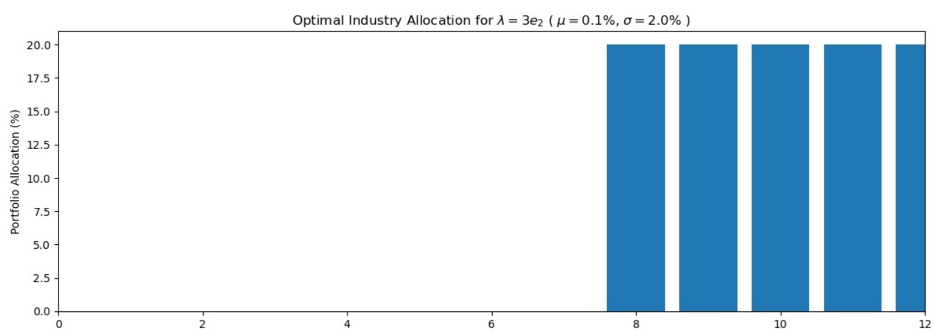

Figure 9: Optimal allocation percentages for industries with a regularization parameter of $3e^{-2}$

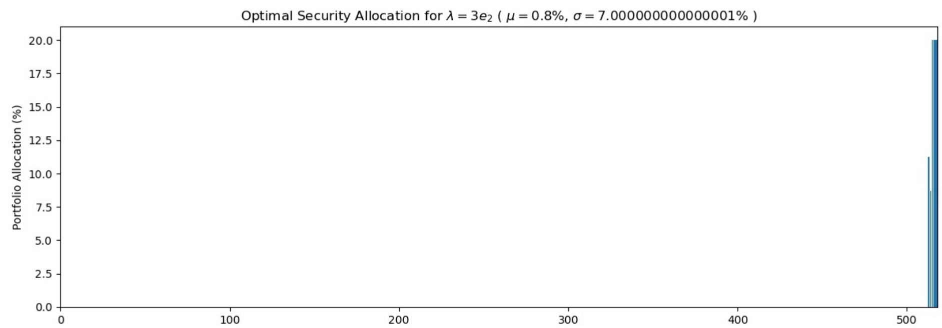

Figure 10: Optimal allocation percentages for securities with a regularization parameter of $3e^{-2}$

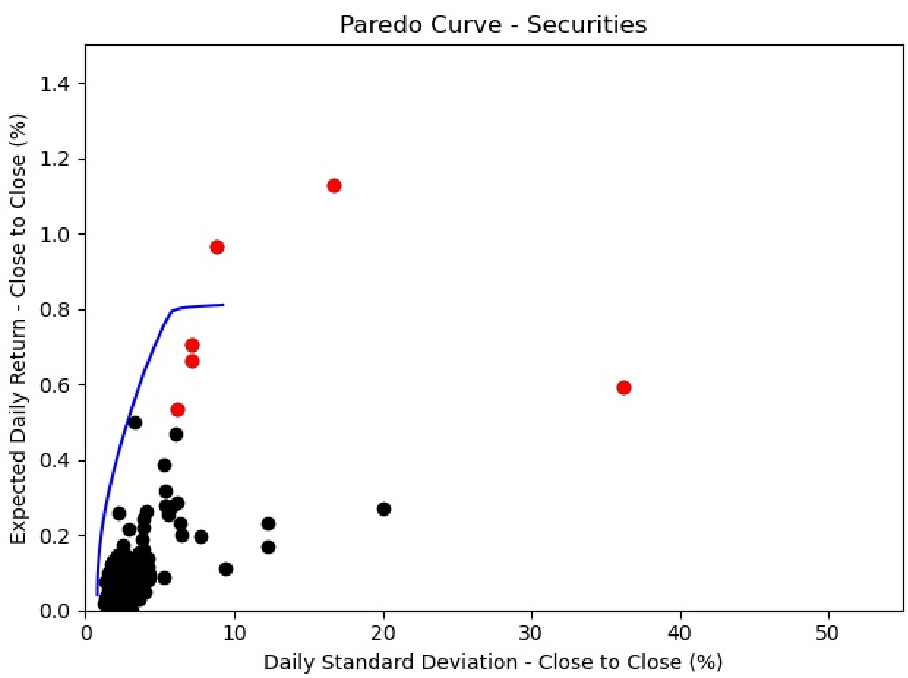

Figure 11: Paredo curve obtained by optimizing over all securities

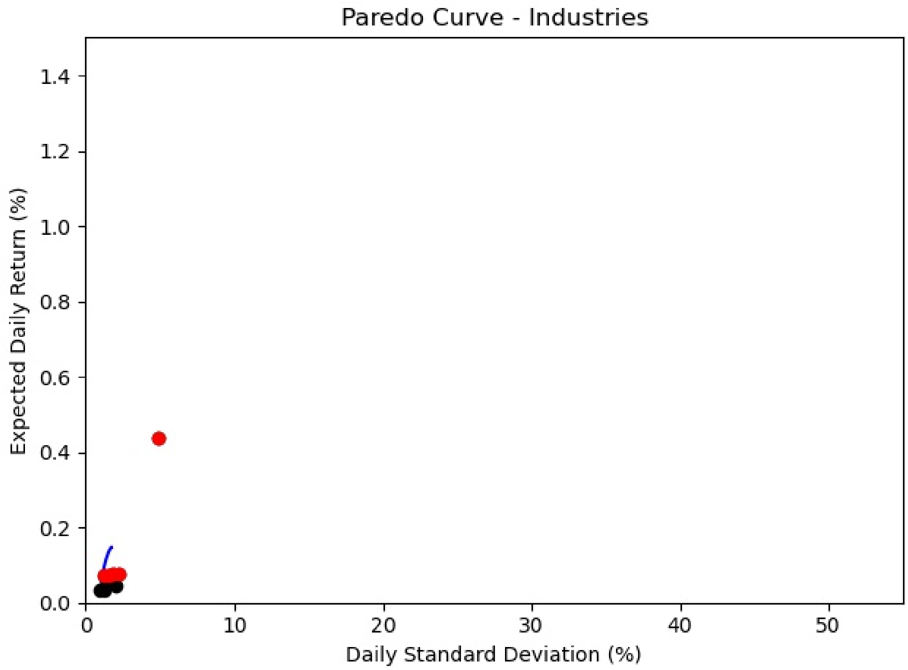

Figure 12: Paredo curve obtained by optimizing over all industries

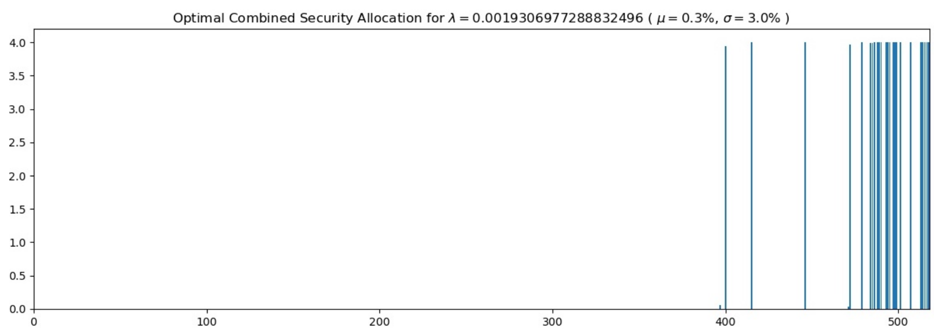

Figure 13: Security allocation percentages of the optimal combined MV portfolio

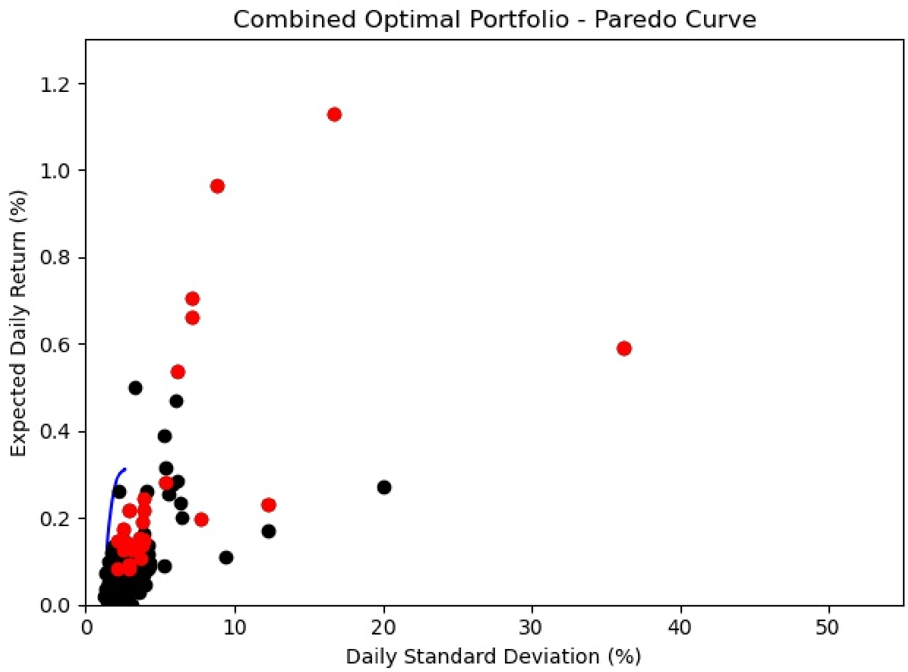

Figure 14: Paredo curve obtained for the combined optimal portfolio

C. Analysis on Conditional Value at Risk Optimization

In this Appendix, we show some metrics and analysis conducted during the Conditional Value at Risk Optimization Process.

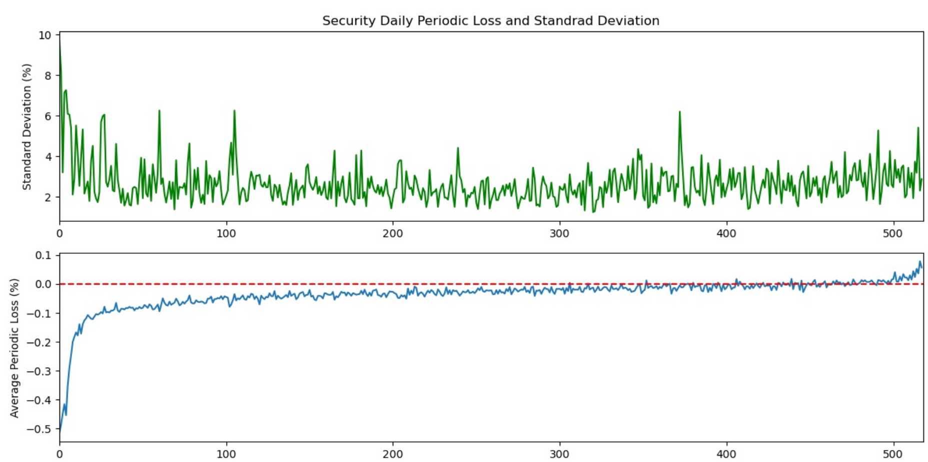

Figure 15: Daily Periodic Loss % and Standard Deviation of all securities, sorted by average Periodic Loss %

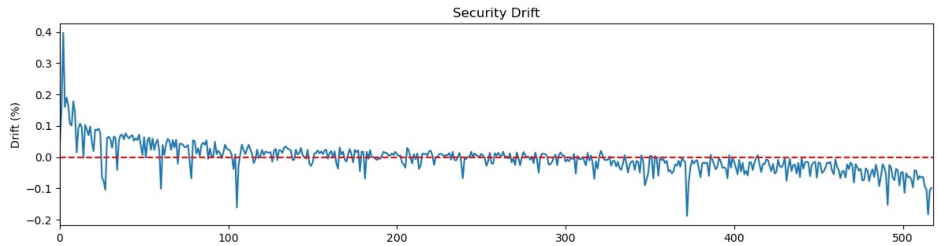

Figure 16: Periodic Drift of all securities, sorted by average security Periodic Loss

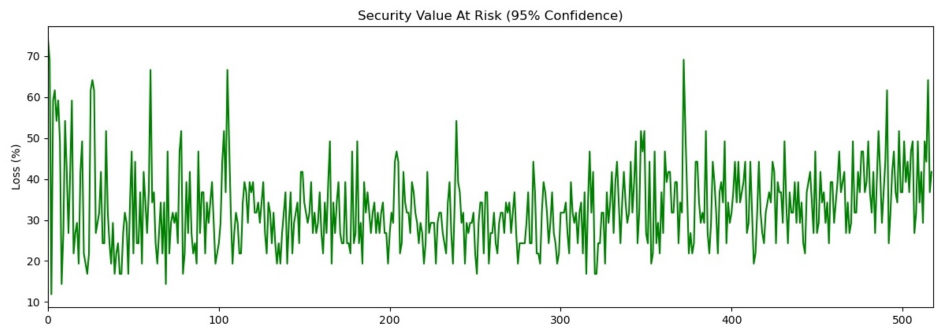

Figure 17: 95%-VaR of all securities simulated from Monte Carlo

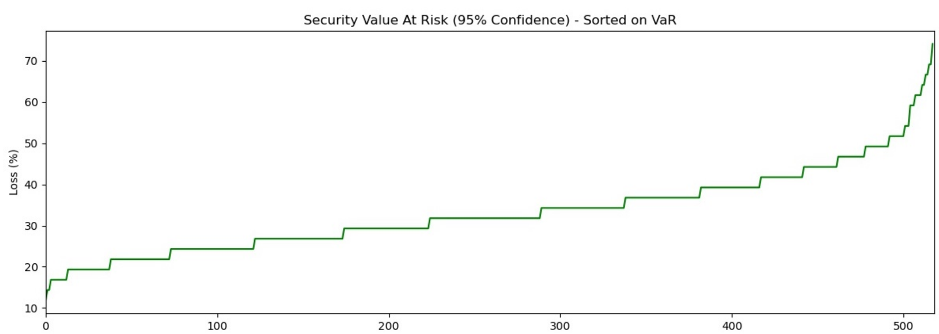

Figure 18: 95%-VaR of all securities simulated from Monte Carlo, sorted by VaR

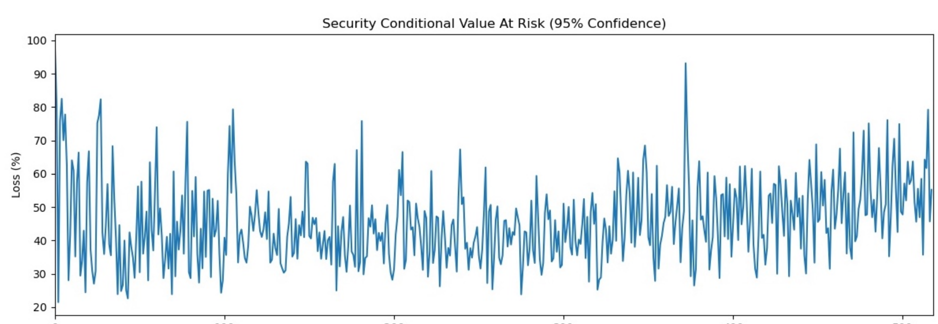

Figure 19: 95%-CVaR of all securities simulated from Monte Carlo

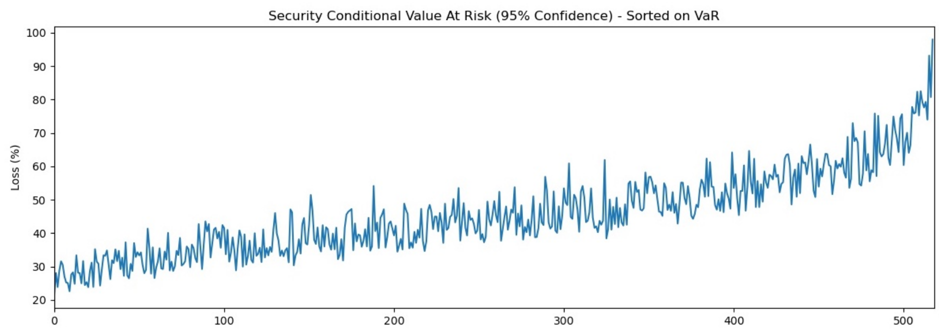

Figure 20: 95%-CVaR of all securities simulated from Monte Carlo, sorted by Var

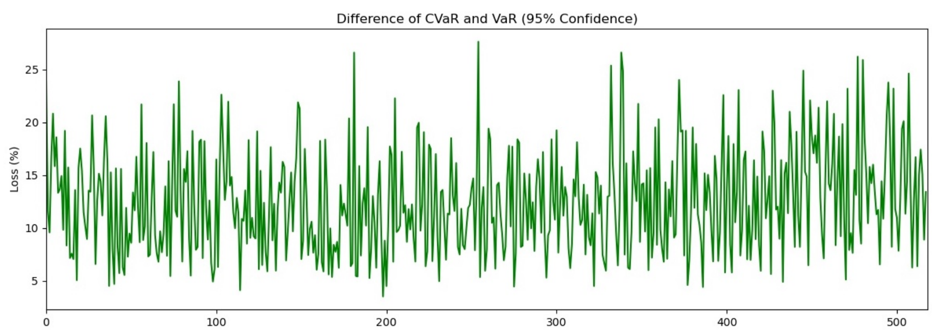

Figure 21: Difference between 95%-VaR and 95%-CVaR of all securities simulated from Monte Carlo

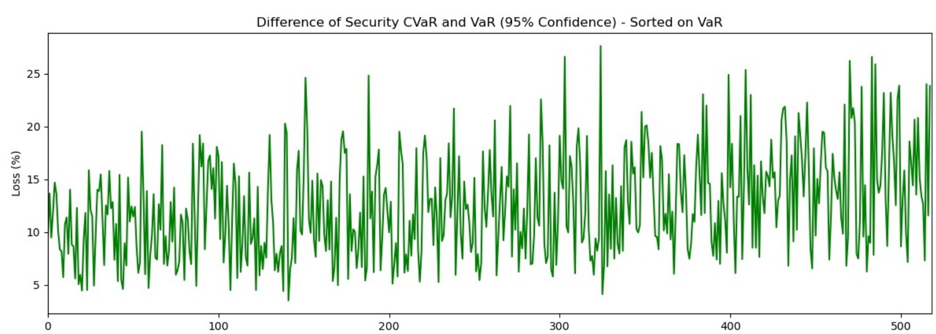

Figure 22: Difference between 95%-VaR and 95%-CVaR of all securities simulated from Monte Carlo, sorted by VaR

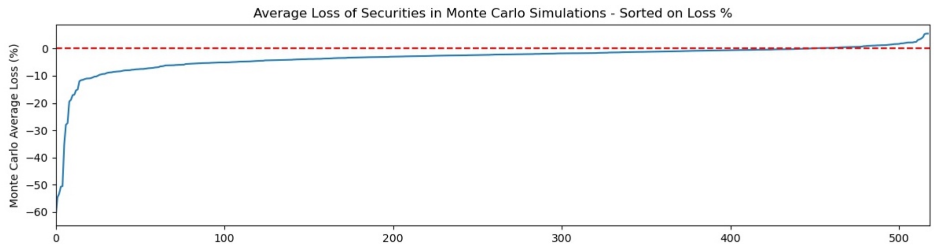

Figure 23: Average Periodic Loss of all securities

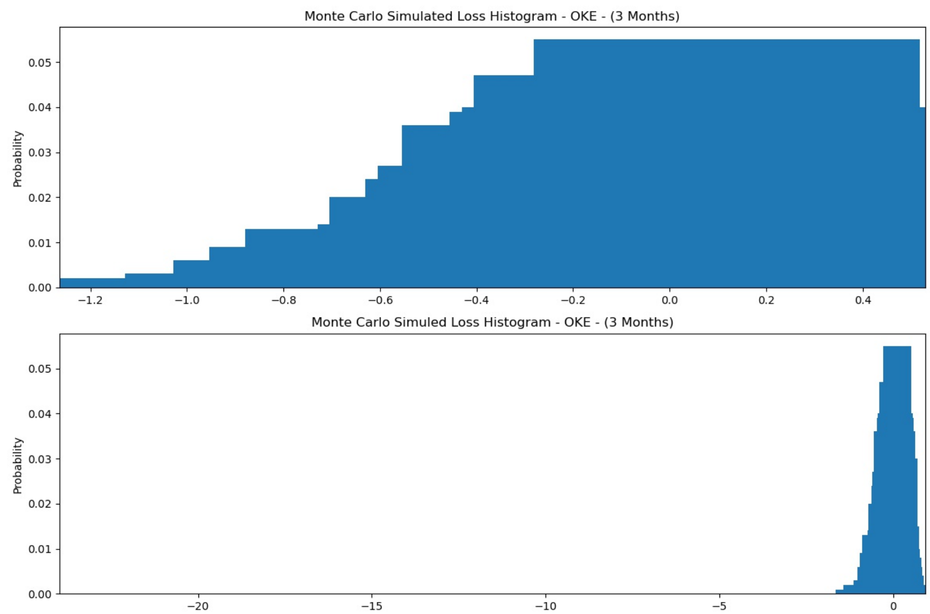

Figure 24: Simulated Loss Distribution for OKE

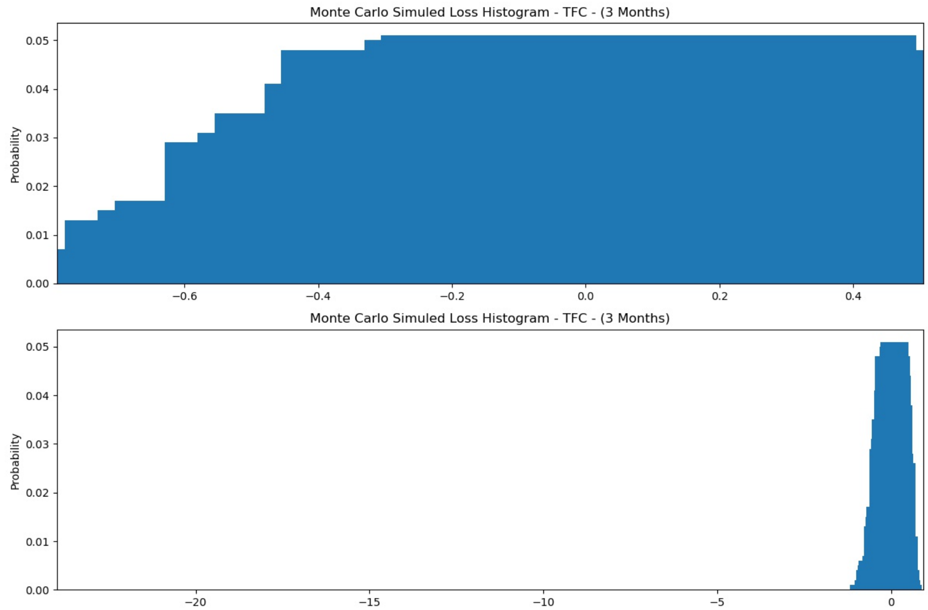

Figure 25: Simulated Loss Distribution for TFC

References

[1] P. Artzner, F. Delbaen, J. M. Eber, and D. Heath, *Thinking coherently*, Risk, 10 (1997), pp. 68–71.

[2] ——, *Coherent measures of risk*, Mathematical Finance, 9 (1999), pp. 203–228.

[3] G. Cornuéjols, J. Peña, and R. Tütüncü, *Optimization methods in finance*, 2nd Edition (2018).

[4] Y. S. Kan and A. Kibzun, *Stochastic programming problems with probability and quantile functions*, Wiley, (1996).

[5] R. Kast, E. Luciano, and L. Peccati, *Var and optimization*, Proceedings of the 2nd International Workshop on Preferences and Decisions, Trento, 1-3 July 1998, (1998).

[6] P. Krokhmal, J. Palmquist, and S. Uryasev, *Portfolio optimization with conditional value-at-risk objective and constraints*, ISE Dept., University of Florida, September, 2001 (Can be downloaded: https://www.ise.ufl.edu/uryasev/files/2011/11/kro_CVaR.pdf), (2001).

[7] H. Mausser and D. Rosen, *Beyond var: From measuring risk to managing risk*, ALGO Research Quarterly, I(2) (1999), pp. 5–20.

[8] J. Morgan, *Riskmetrics technical manual*, JP Morgan, New York, 4th edn. (1996).

[9] G. Pflug, *Some remarks on the value-at-risk and the conditional value-at-risk*, Kluwer Academic Publishers, In: S. Uryasev (Ed) Probabilistic Constrained Optimization: Methodology and Applications. (2000).

[10] R. T. Rockafellar, *Optimization of conditional value at risk*, Princeton Mathematics, Princeton Univ. Press, Vol. 28 (1970).

[11] R. T. Rockafellar and S. Uryasev, *Optimization of conditional value at risk*, Journal of Risk, 2/Number 3 (2000), pp. 21–41.

[12] ——, *Conditional value-at-risk for general loss distributions.*, Research Report 2001-5. ISE Dept., University of Florida, April, 2001 (Can be downloaded: http://www.ise.ufl.edu/uryasev/cvar2.pdf), (2001).

[13] N. Z. Shor, *Minimization methods for non-differentiable functions*, Springer, (1985).

[14] S. Uryasev, *Derivatives of probability functions and some applications*, Annals of Operations Research, 56 (1995), pp. 287–311.