Incorporating Value-at-Risk in Portfolio Selection: An Evolutionary Approach

Chueh-Yung Tsao[*1] **Chao-Kung Liu** [2]

[1]Dept. of Business Administration, Chang Gung University
[2]Dept. of Information Management, Chang Gung University

Abstract

The mean-variance framework for portfolio selection should be revised when investor's concern is the downside risk. This is especially true when the asset returns are not normal. In this paper, we incorporate value-at-risk (VaR) in portfolio selection and the mean-VaR framework is proposed. Due to the two-objective optimization problem faced by the mean-VaR framework, an evolutionary multi-objective approach is applied to construct the mean-VaR efficient frontier. In particular, the NSGA-II is considered here. From the empirical analysis it is found that the risk-averse investor might inefficiently allocate his wealth if his decision is based on the mean-variance framework.

Keywords: NSGA-II, mean-variance efficient frontier, mean-VaR efficient frontier, portfolio selection.

1. Introduction

Portfolio selection is concerned with the problem of optimally allocating the wealth to the riskless and risky assets. For a risk-averse investor, he would like to find the most desirable categories of assets to hold, such that obtaining the highest return and lowest risk. While employing the mean of portfolio return as the measure of return and the variance or standard deviation as the measure of risk, the problem would be transferred into the mean-variance formulation under the assumption of the risky assets returns following multivariate normal distribution or the utility function of the investor being quadratic. A typical proceeding of this analysis starts from constructing a well-known risk-return efficient frontier. Each point on the frontier represents a portfolio of assets. Given the risk preference of the investor, the optimal portfolio is determined by choosing the pair of mean and standard deviation (variance) on the frontier with the best utility and then tracking the weights of the assets forming that portfolio. The analysis mentioned above is the famous Markowitz [8] portfolio theory.

Actually, 'risk' is a conceptual idea and there is no clear mathematical definition of it. In general, the risk of a portfolio comes from the uncertainty of the underlying assets. For a typical investor in the financial market, the risk is the likelihood of losing all or most of a portfolio's value. If that is the case, what the investor concerned with is the 'downside risk' of the assets. The downside risk is the likelihood that a security or other investment will decline in price, or the amount of loss that could result from that potential decline. In the mean-variance formulation, the standard deviation (variance) is used for measuring risk. However, minimizing standard deviation (variance) is equivalent to minimizing downside risk only when the uncertainty of the asset is symmetric, such as the normal random disturbance. Consider an extreme situation where a risky asset never decreases its value. Then the variability of the asset seems not to be dislike.

Value-at-Risk (VaR) is one of the most commonly used methods for measuring downside risk. VaR is defined as the value that can be expected to be lost during severe, adverse market fluctuations. The concept and use of VaR is relatively recent. The use of VaR by major financial firms traces back to the late 1980s to measure the risks of their trading portfolios. Since then, the use of VaR has exploded. The widely used of VaR now by other financial institutions, nonfinancial corporations, and institution investors attributes to the J.P. Morgan's attempt to establish a market standard through its RiskMetrics™ system in 1994 [9]. The Basle Committee on Banking Supervision [1] permitting banks to calculate their capital requirements for market risk using their own proprietary VaR models further increases the growth of VaR.

In this paper, we incorporate the VaR in portfolio selection. In particular, the elitist non-dominated sorting GA (NSGA-II, [2][3]) is applied to sketch the efficient frontier under the mean-VaR

[*] Corresponding Author. Email: cytsao@mail.cgu.edu.tw

framework. From the empirical analysis it is found that the portfolios on the mean-VaR and mean-variance efficient frontiers are equivalent only when we assume that the assets returns follow normal distributions. In addition, the portfolios on the mean-VaR efficient frontier are different from those on the mean-variance efficient frontier when requiring high level VaR and the nonparametric method is used for estimating VaR, which means that the risk-averse investor might inefficiently allocate his wealth if his decision is based on the mean-variance framework.

The rest of this paper is organized as follows. In Section 2, we describe the mean-variance efficient frontier and the mean-VaR efficient frontier. Three commonly used methods for estimating VaR are also reviewed. The NSGA-II is briefly introduced in Section 3, following by the experiment results in Section 4. We conclude this paper in Section 5.

2. Efficient Frontier

Two efficient frontiers are considered in this paper, namely, the mean-variance efficient frontier and the mean-VaR efficient frontier. Consider a portfolio $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)^T$ consisting of n risky assets. Let

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)^T$$

and

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & & \\ \vdots & & \ddots & \vdots \\ \sigma_{n1} & & \cdots & \sigma_{nn} \end{pmatrix}$$

denote the expected return and the variance-covariance matrix of these assets. The expected return and the variance of the portfolio is $\boldsymbol{w}^T\boldsymbol{\mu}$ and $\boldsymbol{w}^T\Sigma\boldsymbol{w}$, respectively. Let $\boldsymbol{r} = (r_1, r_2, \ldots, r_n)^T$ be a vector of random variable representing the future return of the risky assets during a period of time. The one dollar $(1-\alpha)$ VaR of a portfolio $\boldsymbol{w}$ for the given period of time is the value such that

$$\alpha = \Pr(\, \boldsymbol{w}^T\boldsymbol{r} < -\mathrm{VaR}\,)$$

for $\alpha \in (0, 1)$. Note that the one dollar $(1-\alpha)$ VaR is nothing more than the minus of the α percentile for the portfolio return distribution.

It should be noted that one only needs the variance-covariance matrix when obtaining a portfolio's variance. However, obtaining VaR requires more information about the portfolio's distribution.

2.1. Mean-Variance Efficient Frontier

The mean-variance efficient frontier can be regarded as the solutions such that

$$\begin{gathered} \max\ \boldsymbol{w}^T\boldsymbol{\mu} \\ \min\ \boldsymbol{w}^T\Sigma\boldsymbol{w} \\ \text{s.t. } \boldsymbol{w}^T\mathbf{1} = 1. \end{gathered}$$

Each point (μ_e, σ_e^2) on the frontier corresponds to an efficient portfolio representing that when investing among n assets, the least variability one would bear is σ_e^2 given that the payoff level is set to be μ_e, or that the largest payoff one can earn is μ_e when the risk exposure is σ_e^2. This introduces the common 'risk-return' tradeoff in the financial markets. Rational risk-averse investor will make investment decision on the frontier when the risky assets returns follow multivariate normal distribution or the utility function of the investor is quadratic.

In order to efficiently solve the two-objective, quadratic optimization problem mentioned above, the problem can be transferred to a single-objective, quadratic optimization problem

$$\begin{aligned} &\min\ \boldsymbol{w}^T\Sigma\boldsymbol{w} \\ &\text{s.t. } \boldsymbol{w}^T\boldsymbol{\mu} = \mu_e \\ &\text{s.t. } \boldsymbol{w}^T\mathbf{1} = 1 \end{aligned} \tag{1}$$

by varying different level of μ_e. Eq. (1) has a closed-form solution and many package, such as Matlab, Gauss, and others, can easily solve this problem.

2.2. Mean-VaR Efficient Frontier

The mean-VaR efficient frontier can be defined as the solutions such that

$$\begin{gathered} \max\ \boldsymbol{w}^T\boldsymbol{\mu} \\ \min\ \mathrm{VaR} \\ \text{s.t. } \boldsymbol{w}^T\mathbf{1} = 1. \end{gathered}$$

Each point (μ_e, VaR_e) on the frontier corresponds to an efficient portfolio representing that when investing among n assets, the least VaR one would bear is VaR_e given that the payoff level is set to be μ_e, or that the largest payoff one can earn is μ_e when the risk exposure is VaR_e.

In the literature, three methods are commonly used for obtaining VaR, namely, the delta-normal method, the historical simulation method, and the Monte Carlo simulation. The delta-normal method is based on the assumption that the risky assets have a multivariate normal distribution. Under this assumption and that the portfolio is a linear combination of these assets, the portfolio return is also normal and the $(1-\alpha)$ VaR, i.e. α percentile, of the portfolio is

$$\mathrm{VaR} = -(\, \boldsymbol{w}^T\boldsymbol{\mu} - z_\alpha\, (\boldsymbol{w}^T\Sigma\boldsymbol{w})^{1/2}\,)$$

where z_α denote the α percentile for a standard normal distribution. The historical simulation method requires relatively few assumptions about the statistical distributions of the underlying market factors. The only assumption of it is that the future returns distribution of the risky assets is exactly the same as their historical realization. Therefore, given a portfolio $\boldsymbol{w}$ and m historical realizations of the risky assets $\{\boldsymbol{r}_i, i = 1, 2, \ldots, m\}$, the $(1-\alpha)$ VaR is the minus of the αm smallest value of the empirical distribution $\{\boldsymbol{w}^T\boldsymbol{r}_i, i = 1, 2, \ldots, m\}$. The Monte Carlo simulation approach is very similar to the historical simulation method. The major dissimilarity is that, instead of using historical realization as the proxy distribution for future return, in Monte Carlo simulation, one chooses statistical distributions (or processes) that are believed to adequately capture or approximate the possible changes in the assets. The next step is to use pseudo-random generator to generate k pairs of hypothetical assets returns, $\{\boldsymbol{r}_i^{MC}, i = 1, 2, \ldots, k\}$. The $(1-\alpha)$ VaR is the minus of the αk smallest value of the hypothetical distribution $\{\boldsymbol{w}^T\boldsymbol{r}_i^{MC}, i = 1, 2, \ldots, k\}$.

To obtain the mean-VaR efficient frontier, the two-objective problem can efficiently be solved by using a traditional optimization technique when the delta-normal method is used. However, for the historical simulation method and the Monte Carlo simulation method, the objective function can not be presented as a quadratic form. Therefore, the gradient-based optimization technique can not be used here. Alternative method is needed to obtain the mean-VaR efficient frontier. In this paper, an evolutionary method is considered due to its elasticity and multi-point search characteristics. In particular, the NSGA-II is applied here.

3. NSGA-II

Deb and his students suggested an elitist non-dominated sorting GA (termed NSGA-II) in 2000 [2][3]. This method not only uses an elite-preservation strategy, but also uses an explicit diversity-preserving mechanism. Hence, NSGA-Ⅱ is applied to sketch the efficient frontier under the mean-VaR framework in this study. Before describing this method step by step in the following, some symbols are defined here. Define $P(t)$ be the parent population of size N and $Q(t)$ be the offspring population of size N in generation t. Let $R(t) = P(t) \cup Q(t)$. The size of $R(t)$ is $2N$.

In general, the NSGA-II involves five steps. First, the parent and offspring populations (P and Q) are initialized and the R is obtained. The initiation method considered in this study is described in Section 3.1. We perform a non-dominated sorting to $R(t)$ and decide the level front : $F(i)$, i=1,2,…, etc. Here, $F(1)$ means the best non-dominated front. we detailed the method in Section 3.2. Second, we use filter model to keep more than a half of good solutions in $R(t)$. We add to $P(t+1)$ from solutions of $F(1)$ in $R(t)$ and solutions of $F(2)$ in $R(t)$, and so on until the size of $P(t+1)$ more than N, we stop this action. Because the last allowed front is being considered, the size of $P(t+1)$ may more than N. Third, we calculate the crowding distance [5,6] for every solutions of $P(t+1)$ and produce the winner sets, $W(t+1)$, by using Crowded Tournament Selection Operator detailed in Section 3.4. We randomly pick two solutions of $P(t+1)$ and compare which is better to pick the winner for N times so the size of $Winner(t+1)$ is N. Fourth, we use crossover and mutation operators to produce a offspring populations $Q(t+1)$ of size N. Then, we let $R(t+1) = P(t+1) \cup Q(t+1)$ and enter next generation. Fifth, after generation finishing, we let $R = P \cup Q$ again. If $F(i)$ of every solutions in R are not equal to 1, we delete them because we only want to keep the best non-dominated sets. Finally, the result is the final portfolio sets.

3.1. Initiation

The commonly used initiation method is the random initiation. For the portfolio application in this study, one picks up random numbers, $\{v_i, i = 1, 2,\ldots, n\}$, uniformly from [0, 1]. The weights of the portfolio $\{w_i, i = 1, 2,\ldots, n\}$ are obtained by normalizing the v_i such that the summation of the weights is 1, i.e. $w_i = v_i/\Sigma v_i$. In this study, the random initiation method is modified by introducing a threshold, $D\in[0, 1]$. If the random number v_i is greater than the threshold D, the v_i is replaced by 0. Therefore, the lower of D makes the weights of portfolio to be 0 more frequently. It is found in this study that the initiation method plays an important role in the experiments and this simple mechanism can efficiently make the population converge to the efficient frontier more quickly and widely.

3.2. Non-Dominated Sorting of a Population

In order to sort solutions by level, we need find each solution's level. Hence, we use a simple method to do this action. First, we need find the best non-dominated solutions and classify them into level 1. Then, to find next level, we temporarily discard the solutions of level 1 and find the non-dominated solutions in the remaining population. This action is continued until all population solutions are classified into a level.

There are many approaches to be suggested for finding the non-dominated sets, but they usually have different computational complexities. To make matters simple, we can use Naïve and Slow approach, Continuously Updated approach or Kung et al.'s Efficient Method. There are two objectives in our problem in this paper and Kung et al.'s Efficient Method is more efficient and more suitable for finding the non-dominated sets in some problems with few objectives than other two approaches [4] so we use Kung et al.'s Efficient Method to find a non-dominated set. This method will be described in

Section 3.3. In addition to use Kung et al.'s Efficient Method, we let it strictly. We do all Kung et al.'s Efficient Method for every objective function in populations. Then, we intersect all non-dominated sets.

3.3. Kung et al.'s Efficient Method

Kung et al.'s Efficient Method is used to find a non-dominated set [7]. This method sorts the population by descending order of first objective function value first and then recursively halves the population into *top (T)* and *bottom (B)* subpopulations. Here, the top subpopulation is better than bottom subpopulation in the first objective function value. Finally, this method checks something whether any solutions in the bottom subpopulation are not dominated by any solutions in the top subpopulation. If the solutions of *B* that are not dominated by any solutions of *T*, these solutions are combined with *T*.

3.4. Crowded Tournament Selection Operator

The crowded comparison operator compares two solutions and returns the winner of the tournament. It compares two solutions by a level $F(i)$ and a local crowding distance di in the population. Based on these two attributes, we can define the crowded tournament selection operator as follows.

A solution i wins a tournament with another solution j if any of the following conditions are true:
1. If solution i has a better rank, that is $F(i)<F(j)$.
2. If they have the same rank but solution i has a better crowding distance than solution j, that is ,$F(i)=F(j)$ and $di > dj$ [4].

4. Experiment Results

4.1. Preliminary Results

The component assets considered in this paper to construct the portfolio are the component stocks of the TSEC Taiwan 50 index because the TSEC Taiwan 50 index is established by choosing the 50 highest market value stocks on the Taiwan Stock Exchange (TSE), which means that the component stocks are the most representative for the TSE. In order to fully collect price data daily from the sample period of 2001/1/2 to 2005/12/30, only 24 stocks are considered in the experiment.[1]

[1] The details for the 24 stocks are available upon request from the authors.

Table 1: Parameter setting in NSGA-II.

Parameter	Value
Chromosome	Binary
Population Size	100
Number of Generations	300
Select Strategy	Tournament
Crossover Type	Double-Point Crossover
Crossover Probability	0.9
Mutation Probability	0.02
Initiation Threshold	0.1

Four efficient frontiers are constructed in this study, namely, the mean-variance efficient frontier (MV), the mean-VaR efficient frontier based on the delta-normal method (MVaR-N), the mean-VaR efficient frontier based on the historical simulation method (MVaR-H), and the mean-VaR efficient frontier based on the Monte Carlo simulation method (MVaR-M). To implement the experiments, the first frontier is obtained by using the traditional optimization technique and the other three frontiers are obtained by using the NSGA-II. The parameter settings in all of the experiments of the NSGA-II are listed in Table 1.

In order to examine the efficiency of the NSGA-II for constructing the efficient frontiers, the mean-variance efficient frontier is also built by using the NSGA-II (MV-GA). Fig. 1 is the result. The real lines and the star symbols in Fig. 1 are the MVs and the MV-GAs, respectively. The effect of introducing the threshold method in the initiation phase is shown in panel (a). Four levels of threshold are considered, namely, 1, 0.5, 0.3, and 0.1. Note that the threshold which is equal to 1 means that there is no action of the threshold, i.e. pure random initiation. In addition, lower initiation threshold will make the portfolios more concentrated on some specific stocks; causes the portfolios having more diversified objective values in the initial population. It is found from Fig. 1 that the lower the initiation threshold is, the wider the MV-GA is after 300 generations. The threshold is set to be 0.1 in all of the experiments.

The convergence of the MV-GA is so fast. The panel (b) of Fig. 1 shows the MV-GAs under 4 different number of generation, namely, 50, 100, 300, and 500. It is found that the MV-GA approximates the MV very well after 300 generations. In this study, the number of generation is set to be 300 in all of the experiments. Based on the successful application of the NSGA-II to the mean-variance framework, we could suspect that the NSGA-II can also be used to the mean-VaR framework successfully and the mean-VaR efficient frontier found by the NSGA-II would be close enough to the true efficient frontier no matter whatever method is used to estimate VaR, i.e. the delta-normal method, the historical simulation method, or the Monte Carlo simulation method.

4.2. Main Results

In order to compare the efficiency of the four frontiers, i.e. the MV, the MVaR-N, the MVaR-H, and the MVaR-M, we reevaluate the four frontiers under each of the four environments. Take the environment of mean-VaR framework based on the delta-normal estimation method as an example. For each of the four frontiers the means and VaRs are obtained where the VaRs are calculated by using the delta-normal method. Then the four frontiers are all depicted on the same figure. On the figure, it is obvious that the MVaR-N is at least as efficient as the other three frontiers since the MVaR-N is directly constructed under that framework. Thus, one point we are interested in is the relative efficiency of the four frontiers under different environments.

Fig. 2 is the four frontiers under different environments and confidence levels of the VaR. Three confidence levels are considered here, namely, 90%, 95%, and 99%. The real line, dash line, thin line, and dot line on Fig. 2 represent the MV, the MVaR-N, the MVaR-H, and the MVaR-M, respectively. From panel (a) of Fig. 2, the four frontiers are very similar under the 90% and 95% levels, which means that seeking low VaR and high return portfolio is equivalent to seeking low variance and high return portfolio no matter what VaR estimation method are used. Similar results appear in panel (b), (c) and (d). However, the story is different under the 99% VaR requirement. From panel (a), (b), and (d), the MV, the MVaR-N, and the MVaR-M still entwine but the MVaR-H becomes less efficient under the 99% level. The situation gets worse when the mean portfolio return is lower. In panel (c) of Fig. 2, on the other hand, it is found that the MVaR-H apparently dominates the other three frontiers. In sum, the investment decision based on the MV, the MVaR-N, and the MVaR-M makes no difference in all of the confidence levels. When increasing the level of VaR, the MVaR-H distinguishes itself from the other three frontiers.

The experiment results detected in this study is intuitive from both theoretical and practical viewpoints. First, minimizing variance is equivalent to minimizing VaR when delta-normal method is used to estimate VaR. In addition, the Monte Carlo simulation method is indifferent from the delta-normal method when normal distribution is assumed to be the random disturbance of the component stocks. Therefore, investment decision based on the mean-variance framework is the same as the decision made based on the mean-VaR framework when the delta-normal method or the Monte Carlo simulation method is used. Second, while from the stock markets the daily returns usually have negative skewness and excess kurtosis, the asymmetry and fat-tail properties of the distribution make the normal assumption can not be hold in practice. The miss-matching of the normal distribution gets worse when the model is used to examine the extreme or rare events. Consequently, the nonparametric estimation method such as the historical simulation method could better reflect these properties and thus get different results from those of the methods based on the normal assumption.

5. Conclusion

The issues we examine in this study are how to incorporate the value-at-risk (VaR) in portfolio selection and what the consequence is when different frameworks and estimation methods are applied. In this paper, the elitist non-dominated sorting GA (NSGA-II) is applied to sketch the efficient frontier under the mean-VaR framework. It is found that the efficient frontier can be constructed more widely and quickly when introducing the threshold in the initiation phase of the NSGA-II. The reason why the threshold method works well is that it makes the initial population more diversified in terms of the objective functions. Three estimation methods for VaR are examined, namely, the delta-normal method, the historical simulation method, and the Monte Carlo simulation method. In empirical analysis, the portfolios on the mean-VaR efficient frontier are different from those on the mean-variance efficient frontier when the confidence level is 99%, which means that the risk-averse investor might inefficiently allocate his wealth if his decision is based on the mean-variance framework.

6. References

[1] Basle Committee on Banking Supervision, “Amendment to the Capital Accord to Incorporate Market Risks,” Basle, Switzerland: Basle Committee on Banking Supervision (January), 1996.

[2] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, “A fast and elitist multi-objective genetic algorithm: NSGA-Ⅱ,” *Technical Report* 200001, Indian Institute of Technology, Kanpur: Kanpur Genetic Algorithms Laboratory (KanGAL), 2000a.

[3] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan, “A Fast Elitist Non-Dominated Sorting Genetic Algorithm for Multi-Objective Optimization: NSGA-II,” *Proceedings of the Parallel Problem Solving from Nature VI* (*PPSN-VI*), pp. 849-858, 2000b.

[4] K. Deb, “Multi-Objective Optimization using Evolutionary Algorithms,” New York: John Wiley & Sons, 2001.

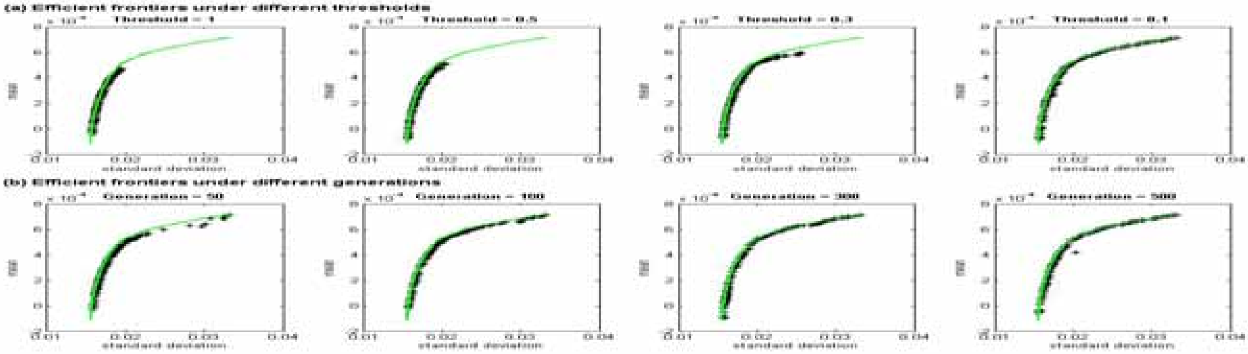

Fig. 1: Mean-variance fficient frontiers under different thresholds and generations.

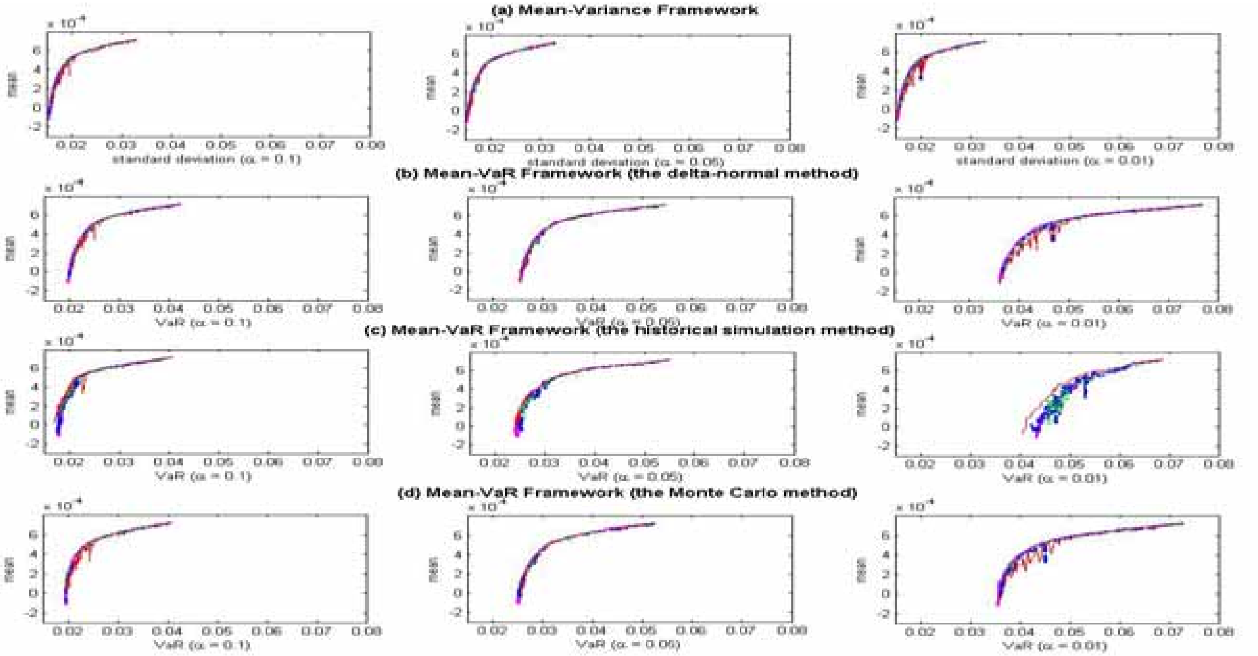

Fig. 2: Efficient frontiers under different envirements and confidence levels.

[5] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation* 6 (2), pp. 182–197, 2002.

[6] M.T. Jensen, "Reducing the Run-Time Complexity of Multiobjective EAs: The NSGA-II and Other Algorithms," *IEEE Transactions on Evolutionary Computation* 7 (5), pp. 503–515, 2003.

[7] H. T. Kung, F. Luccio, and F. P. Preparata, "On Finding the Maxima of a Set of Vectors," *Journal of the Association for Computing Machinery* 22(4), pp. 469-476, 1975.

[8] H. Markowitz, "Portfolio Selection," *Journal of Finance*, Vol. 7, pp. 77-91, 1952.

[9] J.P. Morgan, "RiskMetrics," *Technical Document*. New York: Morgan Guaranty Trust Company (October), 1994.