Evolutionary Optimization for Multiobjective Portfolio Selection Under Markowitz's Model with Application to the Caracas Stock Exchange

Feijoo Colomine Duran[1], Carlos Cotta[2], and Antonio J. Fernández[2]

Abstract Several problems in the area of financial optimization can be naturally dealt with optimization techniques under multiobjective approaches, followed by a decision-making procedure on the resulting efficient solutions. The problem of portfolio optimization is one of them. This chapter studies the use of evolutionary multiobjective techniques to solve such problems, focusing on Venezuelan market mutual funds between years 1994 and 2002. We perform a comparison of different evolutionary multiobjective approaches, namely NSGA-II, SPEA2, and IBEA, and show how these algorithms provide different optimization profiles. The subsequent step of solution selection is done using Sharpe's index as a measure of risk premium. We firstly show that NSGA-II provides similar results to SPEA2 on mixed and fixed funds, and better (according to Sharpe's index) solutions than SPEA2 on variable funds, indicating that NSGA-II provides a better coverage of the region containing interesting solutions for Sharpe's index. Furthermore, IBEA outperforms both NSGA-II and SPEA2 in terms of index value attained. Finally, we also show that this procedure results in a more profitable solution than an indexed portfolio by the Caracas Stock Exchange.

1 Introduction

Finance is a branch of Economics that studies the flow of money and other assets, their acquisition and management by a company, individual or state, and the markets in which they are traded. In other words, it comprises studies concerning the collection and management of money and other valuables such

Universidad Nacional Experimental del Táchira (UNET) Laboratorio de Computación de Alto Rendimiento (LCAR), San Cristóbal, Venezuela, e-mail: fcolomin@unet.edu.ve
· Dept. Lenguajes y Ciencias de la Computación, ETSI Informática, University of Málaga, Campus de Teatinos, 29071 - Málaga, Spain, e-mail: {ccottap,afdez}@lcc.uma.es

as securities, bonds, etc. One of the main challenges for the administration of financial resources is to maintain the profitability and liquidity at times in which the simple act of leaving the money deposited in a bank makes it lose value.

This work focuses on the study of the components that promote an acceptable (above inflation) economic return, as well as the need to obtain better risk diversification as determined by the degree of abhorrence of each investor [15]. Plainly speaking, diversifying amounts using a mechanism *by which all eggs are not put in one basket*, that is, investing in a range of financial sectors whose economic activities result in some benefit and whose economic cycles behave differently from each other. In this context, the risk diversification is achieved by creating a portfolio of investments in several of these financial instruments or sectors.

The area of financial management encompasses a number of theoretical elements and field studies regarding the risk/performance relationship. There is no static optimal solution, and the best portfolio always depends on market evolution. In very general terms, this implies that simultaneous risk minimization and performance maximization are the obvious desired goals. Needless to say, these goals are partially opposed to each other. Several proposals can be found in the literature in this regard. For example, Markowitz's model [18] has become an essential theoretical reference for portfolio selection. However, its practical application has not been as broad as it could, mostly due to the complexity of the method: on one hand, being a quadratic parameterized model its resolution is not trivial; on the other hand, the number of variables involved is high.

The topic addressed by Markowitz relates to the selection of investments, namely the problem of allocating resources among the various options available for that purpose. Prior to the popularization of Markowitz's approach, investment selection involved a costly process of collecting and processing a wide range of information about the companies issuing the assets (primarily shares). This information included, among other things, balance sheets and financial statements, status of the company within the industry and within the market as a whole, the quality of company management, dividend policy, and so on. Markowitz's approach significantly simplified the selection problem by considering asset performance as a stochastic process, focusing solely on the historical log of returns of the issuing companies, and more precisely on three statistical measures of these data: mean, variance and covariance of return rates.

Markowitz developed the model based on the rational behavior of the investor. In other words, the investor wants to maximize her profit and rejects the risk. Therefore, a portfolio will be efficient for her if it provides the highest possible return for a given risk, or equivalently, if it presents the least possible risk for a given level of profitability. The collection of portfolios offering such a combination of risk/profitability is termed the *efficient frontier*, and once known the investor can select her optimal portfolio according to her preferences.

If no additional considerations are made and a specific risk/profitability profile is known, the optimization problem can be solved using quadratic programming. However, this is not usually the case. On one hand, several constraints such as cardinality constraints (i.e., a limit on the number of different investments in the portfolio) or minimum transaction lots can be considered, thus making quadratic programming or other exact techniques infeasible. On the other hand, if no profitability target is fixed a priori (or if a more general investment strategy is sought) the task of finding (or approximating as much as possible) the whole efficient frontier in an efficient way requires the use of powerful optimization techniques. In this scenario the use of metaheuristic techniques is the general norm [3]. These techniques cannot provide optimality proofs for the solutions they obtain, but adequately crafted, they will likely provide optimal or near-optimal solutions to a wide range of continuous and combinatorial optimization problems.

We consider the particular case of nature-inspired metaheuristics, or more precisely, evolutionary algorithms. Not only do these techniques hold an impressive successive-record on different hard optimization tasks; they have also been shown to be extremely effective in solving multiobjective optimization problems. As such, they are quite appropriate to deal with the combined risk/performance optimization. We consider several state-of-the-art second generation approaches for evolutionary multiobjective optimization, and compare them on the basis of sound performance metrics defined in the literature. We also address the subsequent selection step: once the efficient frontier has been identified, there remains the problem of selecting one particular solution according to the risk profile determined by the investor. This latter approach is considered here, and as it will be shown, using Sharpe's index as a guiding measure we are able to identify solutions better than those currently used in indexed portfolios by the Caracas Stock Exchange, a Latin American exchange operating in Venezuela.

2 Background

The work presented in this chapter deals with real investments that are conditioned by two main parameters: (i) profitability, i.e., the returns on the investment, and (ii) risk, i.e., the chances of low (or even negative) returns. Obviously, profitability is a positive element for the investor whereas risk is a negative one. This means that an investor wishes to maximize profitability and minimize risk [1]. This will be formalized within Markowitz's model in

[1] Other parameters such as liquidity or political control of a company might be considered as well.

Section 2.2. Before that, a brief overview of mutual funds will be provided first in Section 2.1.

2.1 Mutual Funds

Mutual funds are instruments that combine the money invested by a group of persons. They are handled by a management office specializing in the administration of investment portfolios, which takes decisions on the purchase of shares, bonds, and other instruments of the market. By combining the money of several investors, mutual funds allow them to participate in larger portfolios than those they could buy individually. There are several types of mutual funds:

1. **Fixed revenue**: The aim of these funds is to invest in state bonds, private bonds, and other instruments that offer a predictable performance if they are maintained up to their expiration. Fixed funds are a way of adjusting to different investment horizons, e.g., they tend to specialize in investments on assets that operate in either a short term, a middle term, or a long term basis. Purchasing assets that participate in fixed funds has diverse advantages such as risk diversification, professional administration, (management), and accessibility (i.e., small investors have access to high investments)
2. **Variable revenue**: These funds try to maximize the profit by investing in shares of companies that quote in the Stock exchange. Because of the very nature of the assets in which they are invested in, these funds have the highest risk associated with them. From a conceptual point of view, the investments shares, if correctly chosen, is the most profitable in the longer term. However, this also implies a higher volatility of the investments (i.e., increased risk).
3. **Mixed revenue**: These funds represent a combination of those funds – fixed and variable revenue– mentioned before. They aim to diversify the investment in stocks of both fixed and variable revenue. Their composition is thus a combination of the different types of assets, and their risk / profitability ratio is intermediate between that of fixed funds and variable funds. The risk obviously depends on the proportion of the investment made is done on the different mutual funds.

Whichever type of fund considered, the investor and/or manager is faced with an optimization problem regarding the composition of the portfolio. One of the most widely used and conspicuous method of addressing this problem is Markowitz's model, described in next section.

2.2 Markowitz's Model and Sharpe's Index

Markowitz's model [18] is a pioneering model in the selection of assets to construct an ideal portfolio. It assumes that the future performance a specific investment can offer can be determined from both experience and investigation. This model is thus applied with the aim of obtaining an optimal portfolio selection. The basic idea is that by analyzing the expected profitabilities of the individual financial assets one can make a correct portfolio selection. Two main components have to be taken into account: profitability and the risk to be assumed by the investor. The investor needs to know the risk level, that is to say, the degree of profitability variation which is measured by the variance defined as:

$$\sigma_i^2(\mathbf{R}) = \sum_{t=1}^{T} \frac{[R_{it} - E(R_i)]^2}{n} \tag{1}$$

where $\mathbf{R} = \{R_{it}\}$, $1 \leqslant i \leqslant n$, $1 \leqslant t \leqslant T$, is a matrix containing the profitability of each asset at each time interval t, $E(R_i)$ is the mean profitability of the i-th asset, and T is the number of intervals in the time horizon. The overall risk of the portfolio is then defined as a weighted quadratic combination of the covariances of the assets included in it, i.e.,

$$\sigma^2(\mathbf{R}|\mathbf{W}) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j \sigma_{ij}(\mathbf{R}) \tag{2}$$

where $\mathbf{W} = \{w_i\}$, $1 \leqslant i \leqslant n$, is a vector comprising the fraction of the budget allocated to each asset ($w_i \geqslant 0$), and $\sigma_{ij}(\mathbf{R})$ is the covariance of the performance of the i-th asset and the j-th asset, defined as:

$$\sigma_{ij}(\mathbf{R}) = \sum_{t=1}^{T} \frac{[R_{it} - E(R_i)]\,[R_{jt} - E(R_j)]}{T} \tag{3}$$

Similar to the risk, the profitability $E(\mathbf{R}|\mathbf{W})$ of a portfolio is defined as the weighted average of the assets involved, i.e.,

$$E(\mathbf{R}|\mathbf{W}) = \sum_{i=1}^{n} w_i E(R_i) \tag{4}$$

Generally speaking, the investor looks for the curve of utility with $E(\mathbf{R}|\mathbf{W}) = \infty$ and $\sigma^2(\mathbf{R}|\mathbf{W}) = 0$, but this not a realistic option as this curve is limited by the existing assets that never have this nature. We note that for the assets without risk (i.e., those with null profit-variance), the utility is equal to the expected profitability because there is no penalization due to the risk.

To evaluate the quality of a portfolio we have to define a measure that accounts for both the profitability and the risk of the assets involved. Such a measure can also allow the comparison between different portfolios. To this end, we have considered Sharpe's index [25], that determines the performance according to the ratio of excess profitability and risk. More precisely,

$$S(\mathbf{R}|\mathbf{W}) = \frac{E(\mathbf{R}|\mathbf{W}) - R_0}{\sigma(\mathbf{R}|\mathbf{W})} \tag{5}$$

where R_0 is the performance of a portfolio without risk. $E(\mathbf{R}|\mathbf{W}) - R_0$ is therefore the excess performance (that is, the extra profit obtained by taking some risks), which is divided by the risk of the portfolio (measured as the standard deviation of returns). Basically, the index indicates how much performance is expected with respect to the risk. The higher the value returned is, the higher the success of the fund management is.

2.3 Related work

An early reference on portfolio optimization with MOEAs is the work of Veradajan *et al.* [29]. They describe the use of NSGA (non-dominated sorting genetic algorithm) [26] to optimize investment portfolios, as an alternative to quadratic programming techniques. In addition to the typical objectives of increasing performance and decreasing risk, a third objective involving the costs of the transactions is also considered. Several variants of the problem involving the presence of additional constraints can be also found in the literature. For example, Chang *et al.* [5] consider limits on the number of assets and their proportion within the portfolio, and use different metaheuristics (tabu search, genetic algorithms, and simulated annealing) to solve the problem. Busetti [4] also consider tabu search and genetic algorithms, in this case for solving the problem with cardinality constraints and transaction costs. Streichert *et al.* [27] deal with a cardinality constrained portfolio selection problem too, using NSGA and evolution strategies. They actually compare different representations of solutions (pure binary, gray binary, and real-valued). Fieldsend *et al.* [11] also deal with this variation of the problem, and more specifically with the case in which the analyst does not know a priori how many instruments should be included in the portfolio, or the degree of risk-performance that can be accepted. They also propose the addition of the cardinality constraints as a third objective to be minimized. Lin *et al.* [17] consider a variant of the problem with fixed transaction costs and minimum transaction lots. They demonstrated that in this case, the selection of the portfolios becomes more complicated because the problem model has to manage mixed integer variables and non-lineal objectives.

From a more general point of view, Mukerjee *et al.* [20] utilize NSGA-II to implement a decision-making multicriteria model used in the risk/performance negotiation by a bank loan manager. Two models with respect to this negotiation were considered. Also in a bank context, Schlottmann and Seese [23] present a survey of different financial applications that can be handled via MOEAs, and encourage the use of specific problem knowledge and hybridization techniques to obtain better algorithms. A more recent perspective on multiobjective evolutionary optimization of portfolios can be found in [21].

3 Material and Methods

Once the problem scenario has been presented, this section is devoted to providing a more precise formulation of the optimization task. Subsequently, the data used in the experiments (corresponding to real market data from a Latin American exchange), as well as the algorithms considered will be described.

3.1 Problem Setting

As stated before, Markowitz's model is based on the assumption that the investor abhors risk, which can be represented as the variability of returns for a certain investment. At the same time, she wants to maximize her profits. Hence, we can consider a portfolio as efficient if it achieves the profit sought by the investor at the minimum risk. The set of efficient portfolios can be calculated by solving the following parametric nonlinear equation:

$$\min \sigma^2(\mathbf{R}|\mathbf{W}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}(\mathbf{R}) \tag{6}$$

subject to:

$$E(\mathbf{R}|\mathbf{W}) = \sum_{i=1}^{n} w_i E(R_i) = V^* \tag{7}$$

$$\sum_{i=1}^{n} w_i = 1 \tag{8}$$

Note that by varying the parameter V^* the optimal solution in each case minimizes the risk of the portfolio for a given target profit. This consideration leads naturally to a multiobjective scenario in which the whole efficient frontier is sought rather than just solving the above equations for different target profits. This efficient frontier comprises Pareto-optimal portfolios, i.e., portfolios whose profitability cannot be increased without increasing the risk as well (and vice versa, the risk cannot be reduced without decreasing the expected return). This bi-objective problem is thus formulated as

$$\min \sigma^2(\mathbf{R}|\mathbf{W}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \sigma_{ij}(\mathbf{R}) \tag{9}$$

$$\max E(\mathbf{R}|\mathbf{W}) = \sum_{i=1}^{n} w_i E(R_i) \tag{10}$$

subject to:

$$\sum_{i=1}^{n} w_i = 1 \tag{11}$$

This basic model corresponds to unconstrained portfolios, in which the investor can allocate any number of investments she desires, and these can be as large or small as wanted. Additional constraints can at any rate be posed on the composition of the portfolio, e.g., cardinality constraints (at most K assets can be included in the portfolio), or size constraints (the fraction of the portfolio allocated to an asset is bounded [2]). We are interested in analyzing carefully the performance of different multiobjective optimizers on the problem, in particular with respect to finding highly desirable solutions according to Sharpe's index. For this reason, we will focus initially on the case of unconstrained portfolios since they provide a more unbiased arena for performance evaluation, and will pave the way for subsequent experimentation with other variants of the problem.

3.2 Data: Venezuelan Mutual Funds

The data used in the experiments is taken from the Caracas Stock Exchange (*Bolsa de Valores de Caracas* - BVC), the only securities exchange operating in Venezuela. More precisely, we have considered data corresponding to the last five years. This time interval is large enough to be representative of the evolution of shares, and not too large to include irrelevant –for prediction purposes- data (the status of funds can fluctuate in the long term, commonly making old data useless for forecasting the future evolution of shares). According to this, our sample – $\sim 35,000$ daily prices of different mutual funds: fixed, variable, and mixed– comprises those funds no older than five years and still available in the BVC [1]. To be precise, we have used weekly market data from year 1994 to year 2002, corresponding to 26 Venezuelan mutual funds: 12 fixed funds, 7 variable funds, and 7 mixed funds. Data up to year 2001 is used for training purposes, whereas data corresponding to the year 2002

[2] Michaud [19] considers that the use of historical data to estimate risk and expected returns introduces an important bias: efficient portfolios can be composed of few, largely uncorrelated assets. Such a portfolio can be unattractive for some investors. However, this problem can be solved by considering constraints on the maximum percentage of the portfolio that a certain asset can represent.

will be used for testing the obtained portfolios with respect to an investment portfolio indexed in the BVC. The relative ratio of share values in successive weeks is calculated to compute the profitability of each fund. This is done for each week in the year, and subsequently averaged to yield the annual weekly mean and thus obtain the annual profit percentage. The covariance matrix of these profitability values is also computed, as a part of Markowitz's model.

3.3 Evolutionary Multiobjective Approaches

The multiobjective portfolio optimization problem posed in this section will be solved via multiobjective evolutionary algorithms (MOEAs). Indeed, multiobjective evolutionary optimization nowadays provides powerful tools for dealing with this kind of problems. A detailed survey of this field is beyond the scope of this work. We refer the reader to [6; 7; 8; 9; 12; 30; 36] among other works for more comprehensive information about this topic. Let us anyway note for the sake of completeness that MOEA approaches can be classically categorized under three major types [36]: (i) aggregation/scalarization, (ii) criterion-based, and (iii) Pareto-dominance based. A fourth class has been defined more recently, namely indicator-based, and will be discussed later. Firstly, let us describe the basis of the three classical approaches.

Aggregation approaches are based on constructing a single scalar value using some function that takes the multiple objective values as input. This is typically done using a linear combination, and the method exhibits several drawbacks, e.g., the difficulty in determining the relative weight of each objective, and the inadequate coverage of the set of efficient solutions, among others. As to the criterion-based approaches, they try to switch priorities between the objectives during different stages of the search (Schaffer's VEGA approach [22] pioneered this line of attack, using each objective to select a fraction of solutions for breeding). This does not constitute a full solution to the problem of approximating the whole efficient front though. Such a solution can be nevertheless obtained via Pareto-based approaches. These are based on the notion of Pareto-dominance. Let f_i, $1 \leqslant i \leqslant n$, represent each of the n objective functions, and let $f_i(x) \prec f_i(y)$ denote that x is better than y according to the i-th objective value. Then, abusing of the notation we use $x \prec y$ to denote that x dominates y when

$$x \prec y \Leftrightarrow [(\exists i : f_i(x) \prec f_j(y)) \wedge (\nexists i : f_i(y) \prec f_j(x))] \qquad (12)$$

The Pareto front (i.e., the efficient front) is therefore the set of non-dominated solutions, i.e., $P = \{x \mid \nexists z : z \prec x\}$. Pareto-based MOEAs use the notion of Pareto-dominance for determining the solutions that will breed and/or the solutions that will be replaced.

In this work we consider three state-of-the-art MOEAs, namely NSGA-II (Non-dominated Sorting Genetic Algorithm II) [10], SPEA2 (Strength Pareto Evolutionary Algorithm 2) [34] and IBEA (Indicator-Based Evolutionary Algorithm) [31]. The first two fall within the Pareto-based class, and are the second-generation version of two previous algorithms –NSGA [26], and SPEA [33] respectively. As such, they rely on the use of elitism (an external archive of non-dominated solutions in the case of SPEA2, and a plus-replacement strategy –keeping the best solutions from the union of parents and offspring– in the case of NSGA-II). More precisely, the central theme in these algorithms is assigning fitness to individuals according to some kind of non-dominated sorting, and preserving diversity among solutions in the non-dominated front. NSGA-II does this by sorting the population in non-domination levels. First of all, the set of non-dominated solutions is extracted from the current population $\mathcal{P}$; let this set be termed $\mathcal{F}_1$, and let $\mathcal{P}_1 = \mathcal{P} \setminus \mathcal{F}_1$. Subsequently, while there exist solutions in $\mathcal{P}_i$, $i \geqslant 1$, a new front $\mathcal{F}_{i+1}$ is extracted, and the procedure repeated. This way, each solution is assigned a rank, depending on the front it belongs to (the lower, the better). Such a rank is used for selection. To be precise, a binary tournament is conducted according to the domination level, and a crowding distance is utilized to break domination ties (thus spreading the front).

As to SPEA2, it uses an external archive of solutions that is used to calculate the "strength" of each individual i (the number of solutions dominated by or equal to i, divided by the population size plus one). Selection tries to minimize –via binary tournaments– the combined strength of all individuals not dominated by competing parents. This fitness calculation is coarse-grained, and may not always be capable of providing adequate guidance information. For this reason, a fine-grained fitness assignment is used, (i) taking into account both the external archive and the current population, and (ii) incorporating a nearest-neighbor density estimation technique (to spread the front). As a final addition with respect to SPEA, a sophisticated archive update strategy is used to preserve boundary conditions (see [34]).

The third algorithm considered is IBEA, which as its name indicates falls within the indicator-based class. Algorithms in this class approach multi-objective optimization as a procedure aimed at maximizing (or minimizing) some performance indicator. Many such indicators are based on the notion of Pareto-dominance and hence, this class of algorithms is in many respects related to these Pareto-based approaches. Nevertheless, it is necessary to note that they deserve separate treatment due to the philosophy behind them. Actually, in some sense, indicator-based algorithms can be regarded as a collective approach, where selective pressure is exerted to maximize the performance of the whole population. Consider, for example, an IBEA approach based on the hypervolume indicator. This indicator provides information on the hypervolume of the fitness space that is dominated by a certain set of solutions. This definition includes singletons (sets of a single solution), and therefore can be used to compare two individuals. This way, it can be used for selection purposes. However when it comes to replacement, a global perspective is used: the solution whose substitution results in the best value of the indicator for the whole population is taken out. In this work, we have considered an IBEA based on the ε-indicator [35].

In all the algorithms considered, solutions, i.e., a vector of rational values in the $[0, 1]$ range indicating the fraction of the portfolio devoted to each fund, are represented as binary strings. Each fund is assigned 10 bits, yielding a raw weight $\bar{w}_i$. These weights are subsequently normalized as $w_i = \bar{w}_i / \sum_j \bar{w}_j$ to obtain the actual composition of the portfolio. Evaluation is done by computing the risk and return of the portfolio using the formulation depicted before. As to reproduction, we consider standard operators such as two-point crossover and bit-flip mutation.

4 Results

The experiments were conducted with the three algorithms described earlier, namely NSGA-II, SPEA2 and IBEA. We have utilized the PISA library (A Platform and Programming Language Independent Interface for Search Algorithms) [2], which provides an implementation of these two algorithms. The crossover rate is $P_x = 0.8$, the mutation rate is $P_m = 1/\ell$, and the population size is 2ℓ, where ℓ is the total number of bits in a solution. The algorithms run for a maximum number of 100 generations. The number of runs per data set is 30.

4.1 Front Analysis

The first part of the experimentation deals with the analysis of the Pareto fronts obtained. The results obtained are graphically depicted in Figs. 1–3. As can be seen, the grand fronts generated by either algorithm seem to be very similar, although the grand front found by IBEA appears to be slightly more spread for fixed funds. To analyze the extent of the significant difference in the performance more carefully we have considered two well-known performance indicators: the hypervolume indicator [32] and the R_2 indicator [13]. As mentioned before, the first one provides an indication of the region in the fitness space that is dominated by the front (and hence the larger, the better). As to the second indicator, it estimates the extent to which a certain front approximates another one (the true Pareto-optimal front if known, or a reference front otherwise). We have considered the unary version of this indicator, taking the combined NSGA-II/SPEA2/IBEA Pareto front as a reference set. Being a measure of distance to the reference set, the lower a R_2 value, the better.

Figs. 4 and 5 show the distribution of these two indicators for the experiments realized. Let us first consider the hypervolume distribution. SPEA2 appears to provide slightly worse values of this indicator with respect to NSGA-II. Actually, NSGA-II is better (with statistical significance at the standard 0.05 level, according to a Wilcoxon ranksum test [16]) on fixed and mixed funds, and provides a negligible difference on variable funds. On the other hand, IBEA exhibits an interesting behavioral pattern with notably better results than both NSGA-II and SPEA2 on fixed funds, no difference on mixed funds, and clearly worse results on variable funds (in all cases, with statistical significance as before). A similar pattern is observed when the R_2 indicator is considered. NSGA-II compares favorably to SPEA2 in all the three types of funds, and IBEA varies from providing the best results on fixed funds to the worst ones on variable funds. Notice that all differences are statistically significant, except SPEA2 vs IBEA on variable funds.

Among the three types of funds, it is clear that the front corresponding to variable funds is the longest one, spreading from very low risk/low profit solutions to high risk/high profit portfolios. On the contrary, the front generated for mixed funds is much more focused on a regime than can be described as low risk/moderate profit. As to fixed funds, they cover a risk spectrum similar to that of variable funds, but the extreme points of attainable profit are well within the range of profit values found for variable funds. A more precise perspective of the particular risk/profit tradeoffs attained by each of

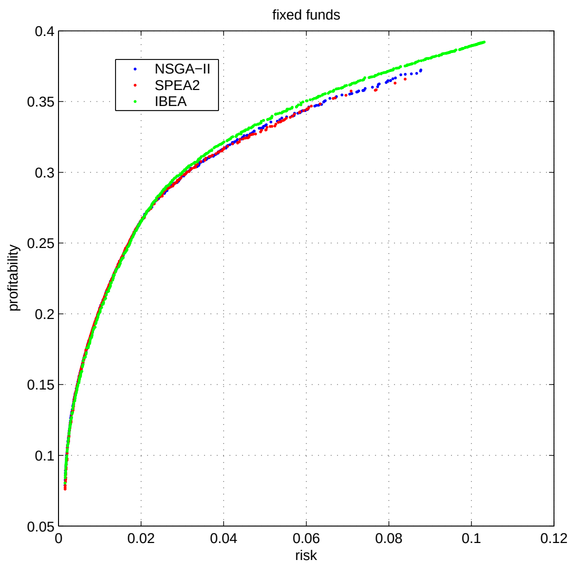

Fig. 1 Comparison of the Pareto fronts found by NSGA-II, SPEA2 and IBEA on fixed funds.

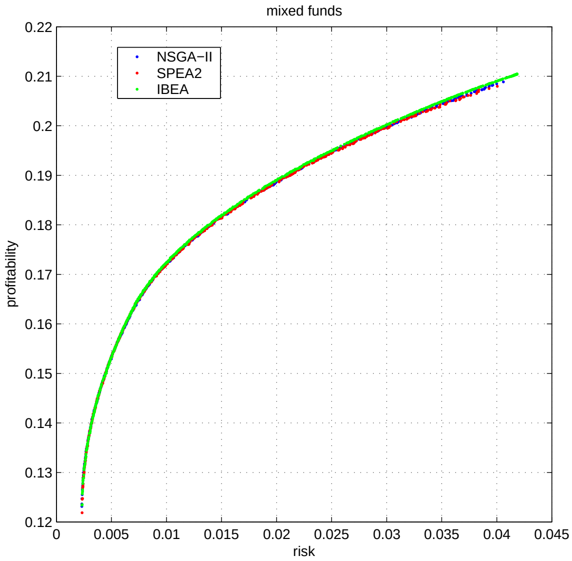

Fig. 2 Comparison of the Pareto fronts found by NSGA-II, SPEA2 and IBEA on mixed funds.

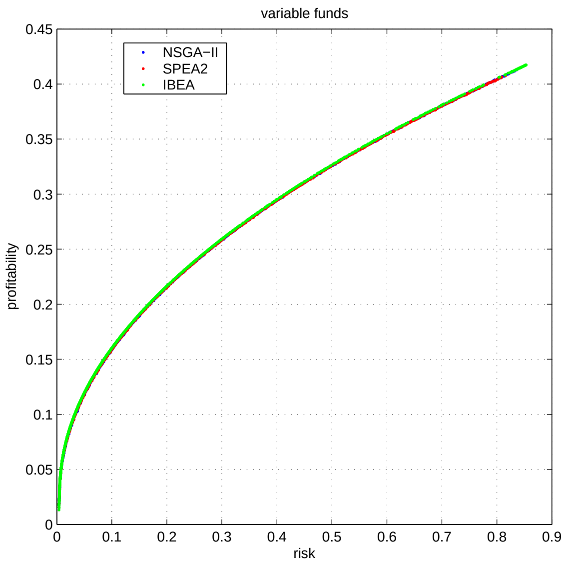

Fig. 3 Comparison of the Pareto fronts found by NSGA-II, SPEA2 and IBEA on variable funds.

the algorithms on the different types of funds will be provided in next section via the use of Sharpe's index.

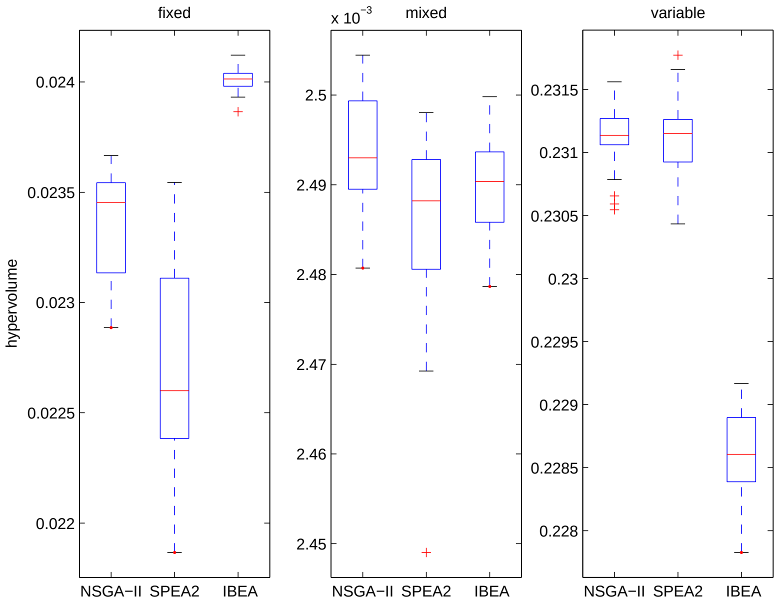

Fig. 4 Boxplot of the hypervolume indicator for NSGA-II, SPEA2 and IBEA.

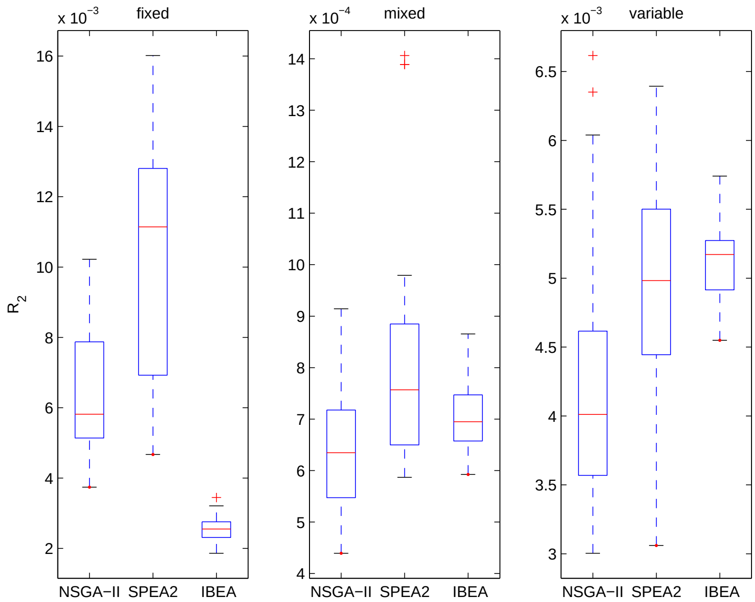

Fig. 5 Boxplot of the R_2 indicator for NSGA-II, SPEA2 and IBEA.

4.2 Use of Sharpe's Index

Sharpe's index has been used for decision-making purposes, enabling the selection of a single solution out of the whole efficient front. Recall that this

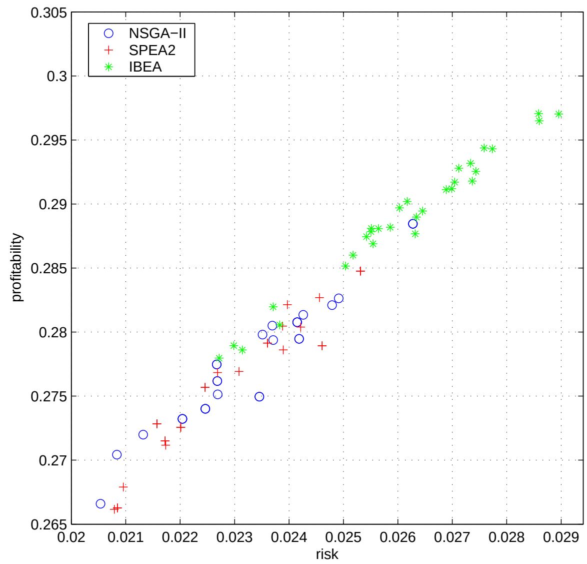

Fig. 6 Best solution (fixed funds) in each run (according to Sharpe's index) found by NSGA-II, SPEA2 and IBEA

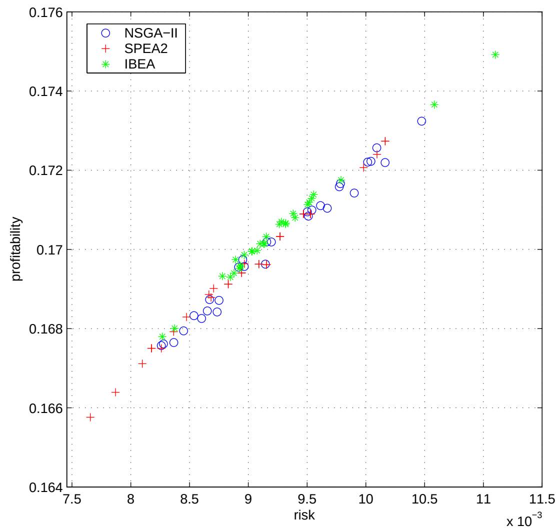

Fig. 7 Best solution (mixed funds) in each run (according to Sharpe's index) found by NSGA-II, SPEA2 and IBEA

index measures how much excess profit per risk unit is attained by a certain portfolio. Depending on the particular shape of the observed front (which

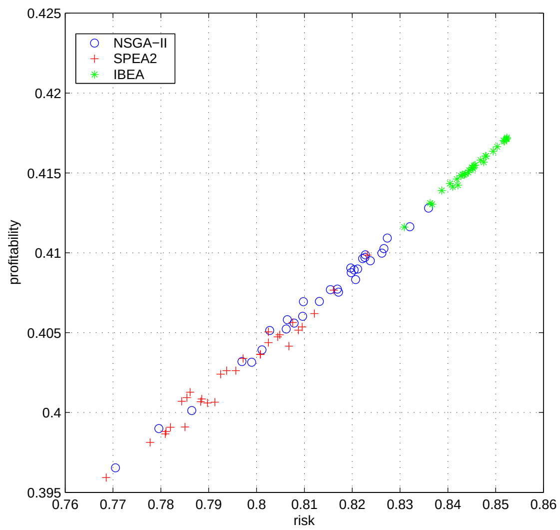

Fig. 8 Best solution (variable funds) in each run (according to Sharpe's index) found by NSGA-II, SPEA2 and IBEA

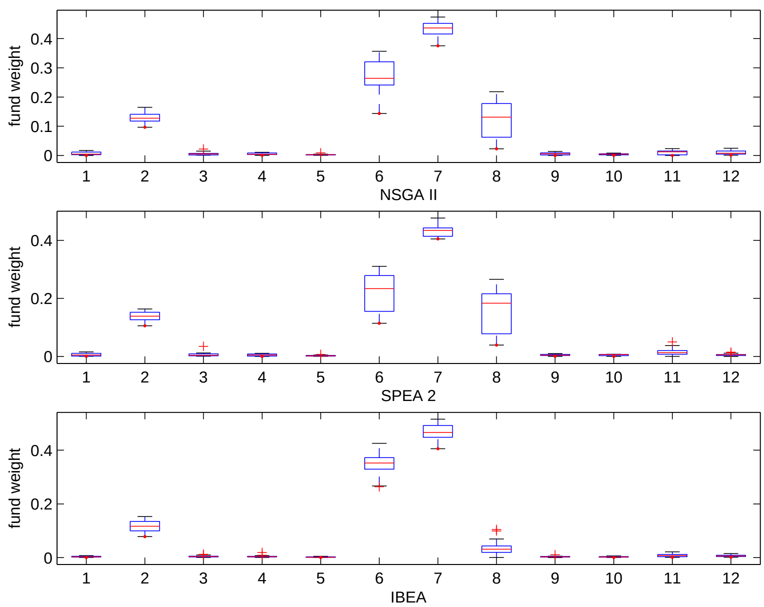

Fig. 9 Portfolio distribution (fixed funds) in solutions selected according to Sharpe's index. (Top) NSGA-II (middle) SPEA2 (bottom) IBEA

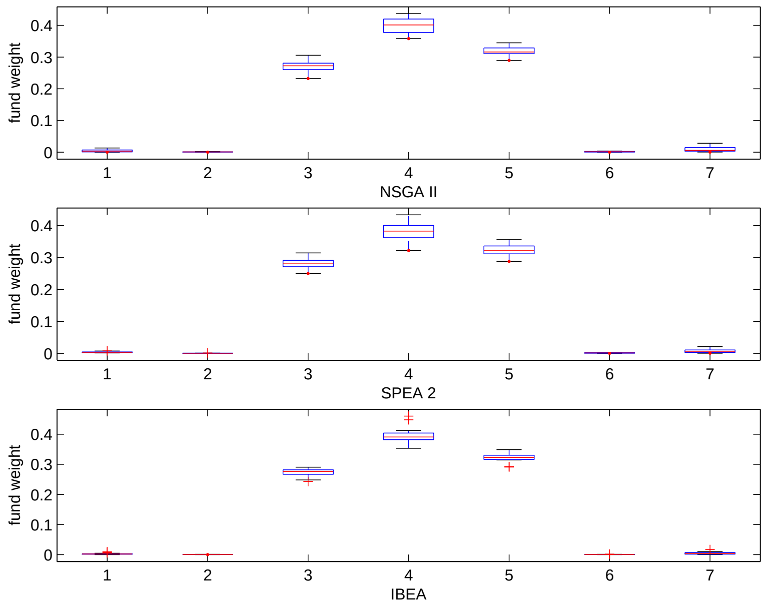

Fig. 10 Portfolio distribution (mixed funds) in solutions selected according to Sharpe's index. (Top) NSGA-II (middle) SPEA2 (bottom) IBEA

depends on the assets that can be potentially included in the portfolio), this solution can correspond to a different risk/profit combinations.

This is illustrated in Figs. 6–8, where the best final solution (according to its Sharpe's index) provided by each algorithm on each of the 30 runs is shown for each type of fund. Best solutions tend to be arranged close to a line whose slope is the optimal value of Sharpe's index. Moreover, solutions are generally clustered in a relatively small range of risk/profit combinations. This indicates all algorithms typically provide solutions with a stable risk/profit profile. Indeed, the composition of portfolios tends to be stable as well, as shown in Figs. 9–10: NSGA-II, SPEA2 and IBEA agree on which funds should be included in the portfolio in each situation, and the variability of percentages (viz. the vertical size of boxes in the boxplot) is small, particularly in variable funds (where investments are mainly concentrated in fund #4, *Mercantil*) and mixed funds (where investments are stably distributed among three funds, *Ceiba*, *Mercantil*, and *Provincial*). In the case of fixed funds there seems to be a higher variability in the percentages of two funds (*Exterior RF* and *Primus RF*) due to their similar profiles.

Another interesting aspect concerns the distribution of Sharpe's index values obtained in each run. Fig. 12 shows a boxplot of Sharpe's index values for the 30 runs of each algorithm on each type of fund. NSGA-II and SPEA2 perform similarly, except on variable funds, where NSGA-II is clearly better. However, IBEA outperforms both NSGA-II and SPEA2 on all types of funds

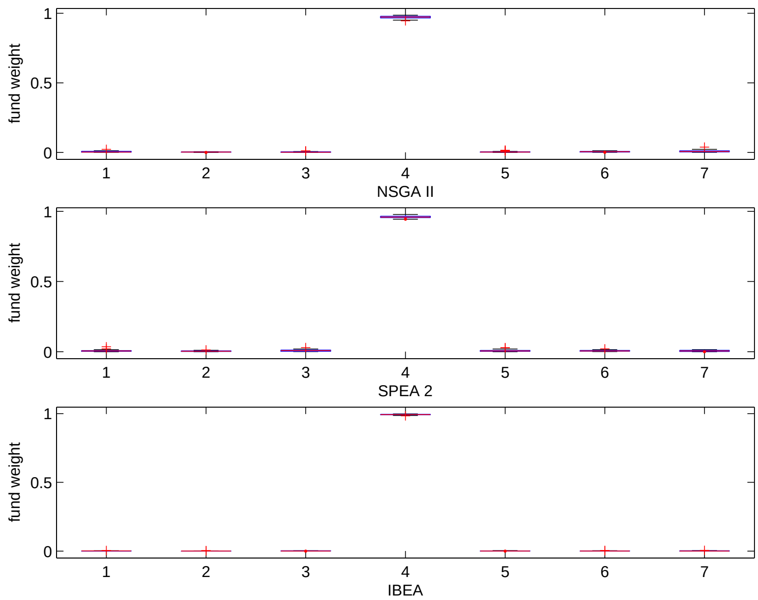

Fig. 11 Portfolio distribution (variable funds) in solutions selected according to Sharpe's index. (Top) NSGA-II (middle) SPEA2 (bottom) IBEA

Table 1 Comparison of the best solutions (according to Sharpe's index) found by NSGA-II, SPEA2 and IBEA.

	Fixed Funds			Mixed Funds			Variable Funds		
	NSGA-II	SPEA2	IBEA	NSGA-II	SPEA2	IBEA	NSGA-II	SPEA2	IBEA
$E(\mathbf{R}\|\mathbf{W})$	.2775	.2821	.2881	.1697	.1690	.1697	.4128	.4099	.4172
$\sigma^2(\mathbf{R}\|\mathbf{W})$	.0227	.0240	.0255	.0090	.0087	.0089	.8359	.8232	.8523
Sharpe's index	1.034	1.036	1.041	.5067	.5060	.5084	.3183	.3175	.3200
$E_{2002}(\mathbf{R}\|\mathbf{W})$	.2367	.2457	.2365	.2455	.2468	.2453	.5392	.5342	.5432

(clear from visual inspection, and further verified by a Wilcoxon ranksum test).

Finally, the best overall solutions found by each of the algorithms are compared to an indexed portfolio in the Caracas Stock Exchange. To this end, we consider data for the year 2002, which was not seen during the optimization process. Table 1 displays the objective values for the best evolved portfolios, and the profit projection for 2002. As a reference, the mentioned indexed portfolio (IBC) achieves a profit of .1988 for 2002. It can thus be seen that the evolved portfolios are notoriously better that this latter portfolio.

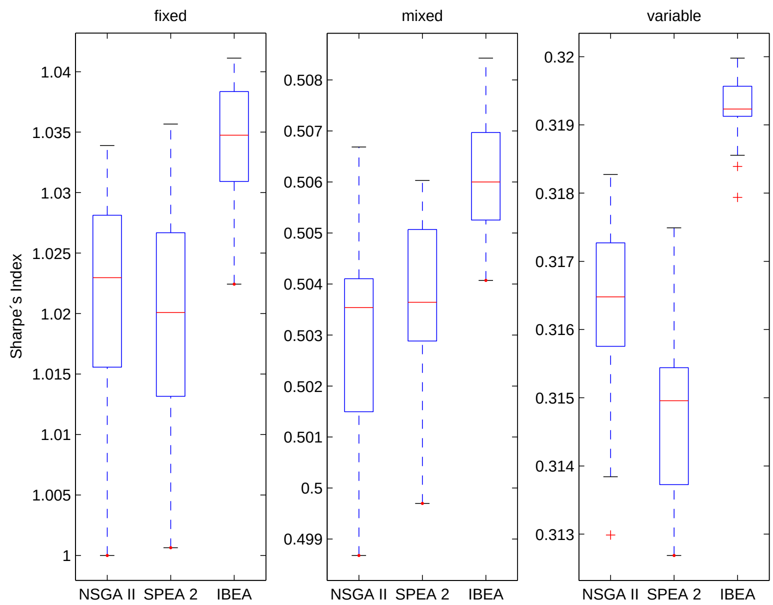

Fig. 12 Boxplots of Sharpe's index values attained by NSGA-II, SPEA2 and IBEA on fixed funds (right), mixed funds (middle) and variable funds (left).

5 Conclusions

Portfolio optimization is a natural arena for multiobjective optimizers. In particular, MOEAs have both the power and the flexibility required to successfully deal with this kind of problems. In this sense, this work has analyzed the performance of three state-of-the-art MOEAs, namely NSGA-II, SPEA2, and IBEA on portfolio optimization, using real-world mutual funds data taken from the Caracas Stock Exchange. Although the algorithms performed similarly from high level –with the exception of fixed funds, where IBEA provides a wider and deeper front– a closer look indicates that they offer different optimization profiles for this problem. NSGA-II is capable of advancing deeper towards some regions of the Pareto front (with statistical significance at the standard 0.05 level in the case of fixed funds and mixed funds), and IBEA lags behind the other two algorithms on variable funds.

Quite interestingly, when the subsequent decision-making step is approached and a single solution is selected from the Pareto front, the comparison turns out to be favorable to IBEA in all the cases. Furthermore, NSGA-II is better than SPEA2 on the problem scenario –variable funds– on which it did not achieve better quality indicators than the latter. More precisely, using Sharpe's index –based on a profit/risk ratio– to identify the best solution from the Pareto front provides significantly better values when using NSGA-II than SPEA2 on variable funds. This indicates a much better coverage of the region where such solutions lie. There is no statistically significant difference in the case of fixed and mixed funds. Likewise, IBEA provides much better solutions in this latter case, even when the quality indicators were worse than those of NSGA-II and SPEA2. This fact illustrates a recurrent theme in multiobjective optimization, i.e., the extent of the usefulness of approximating the whole Pareto front in practical problem scenarios. The fact that a deeper, wider, and more complete the Pareto front returned by an algorithm is better for any problem is based on a reasonable premise: providing the best set of solutions for the decision-maker to make the final selection. However, in some situations the details of how this decision-maker makes this decision cannot be ignored when evaluating the multiobjective optimizer. In other words, the best set of solutions is not necessarily the largest or the most diverse set, but the set that achieves a better coverage of the region in the search space thet the decision-maker is going to prefer. Portfolio optimization under Markowitz's model using Sharpe's index for selection is a good example of this situation.

Future work will be directed at analyzing other variants of the problem where additional constraints are introduced, e.g., cardinality constraints, minimum/maximum percentage of assets, etc. This analysis will pave the way for the development of ad hoc MOEAs, where we plan to integrate specific knowledge on the problem and on the subsequent decision-making procedure. Another line of future research concerns the measure of risk. While we have focused on variance here, this is by no means the unique available option. As an alternative, we may for example consider value at risk, i.e., the maximum loss that can take place at a certain confidence level. A related measure is the conditional value at risk, namely the expected shortfall in the worst q% of cases, where q is a parameter. Other possible measures are Jensen index [14], Treynor index [28], or models emanating from capital asset pricing theory (CAPM) [24], among others. An analysis of these alternatives is underway.

Acknowledgements

This work is partially supported by project TIN2008-05941 of the Spanish Ministry of Science and Innovation.

References

1. AVAF (2003) Annual report of the venezuelan fund management association. `http://www.avaf.org/`
2. Bleuler S, Laumanns M, Thiele L, Zitzler E (2003) PISA—a platform and programming language independent interface for search algorithms. In: Fonseca CM, et al (eds)

Evolutionary Multi-Criterion Optimization 2003, Springer-Verlag, Berlin Heidelberg, Lecture Notes in Computer Science, vol 2632, pp 494–508
3. Blum C, Roli A (2003) Metaheuristics in combinatorial optimization: Overview and conceptual comparison. ACM Comput Surv 35:268–308
4. Bussetti FR (2000) Metaheuristic approaches to realistic portfolio optimisation. Master's thesis, University Of South Africa
5. Chang TJ, Meade N, Beasley J, Sharaiha Y (2000) Heuristics for cardinality constrained portfolio optimisation. Comput Oper Res 27:1271–1302
6. Coello Coello C, Van Veldhuizen DA, Lamont GB (2002) Evolutionary Algorithms for Solving Multi-Objective Problems, Genetic Algorithms and Evolutionary Computation, vol 5. Kluwer Academic Publishers
7. Coello Coello CA (2006) 20 years of evolutionary multi-objective optimization: What has been done and what remains to be done. In: Yen GY, Fogel DB (eds) Computational Intelligence: Principles and Practice, IEEE Computational Intelligence Society, pp 73–88
8. Coello Coello CA, Lamont GB (2004) Applications of Multi-Objective Evolutionary Algorithms. World Scientific, New York
9. Deb K (2001) Multi-Objective Optimization Using Evolutionary Algorithms. John Wiley & Sons, Chichester, UK
10. Deb K, Agrawal S, Pratab A, Meyarivan T (2000) A fast elitist non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. In: Schoenauer M, et al (eds) Parallel Problem Solving from Nature VI, Springer-Verlag, Berlin Heidelberg, Lecture Notes in Computer Science, vol 1917, pp 849–858
11. Fieldsend J, Matatko J, Peng M (2004) Cardinality constrained portfolio optimisation. In: Yang Z, et al (eds) Fifth International Conference on Intelligent Data Engineering and Automated Learning, Springer-Verlag, Berlin Heidelberg, Lecture Notes in Computer Science, vol 3177, pp 788–793
12. Fonseca CM, Fleming PJ (1993) Genetic algorithms for multiobjective optimization: Formulation, discussion and generalization. In: Forrest S (ed) Fifth International Conference on Genetic Algorithms, University of Illinois at Urbana-Champaign, Morgan Kauffman Publishers, San Mateo, California, pp 416–423
13. Hansen M, Jaszkiewicz A (1998) Evaluating the quality of approximations to the nondominated set. Tech. Rep. IMM-REP-1998-7, Institute of Mathematical Modelling Technical University of Denmark
14. Jensen MC (1968) The performance of mutual funds in the period 1945 - 1964. J Financ 23:383–417
15. Knight F (1921) Risk, Uncertainty, and Profit. Hart, Schaffner & Marx; Houghton Mifflin Company
16. Lehmann E, D'Abrera H (1998) Nonparametrics: Statistical Methods Based on Ranks. Prentice-Hall, Englewood Cliffs, NJ
17. Lin D, Li X, Li M (2005) A genetic algorithm for solving portfolio optimization problems with transaction costs and minimum transaction lots. In: Advances in Natural Computation, Springer-Verlag, vol 3612, pp 808–811
18. Markowitz HM (1952) Portfolio selection. J Financ 7:77–91
19. Michaud R (1989) The markowitz optimization enigma: Is optimized optimal? J Financ 45(1):31–42
20. Mukerjee A, Biswas R, Deb K, Mathur AP (2002) Multiobjective evolutionary algorithms for the risk-return trade-off in bank loan management. Int T Oper Res 9:583–597
21. Radhakrishnan A (2007) Evolutionary algorithms for multiobjective optimization with applications in portfolio optimization. Master's thesis, North Carolina State University
22. Schaffer JD (1985) Multiple objective optimization with vector evaluated geneticalgorithms. in genetic algorithms and their applications. In: Grefenstette JJ (ed) First International Conference on Genetic Algorithms, L. Erlbaum Associates, pp 93–100

23. Schlottmann F, Seese D (2002) Hybrid multi-objective evolutionary computation of constrained downside risk-return efficient sets for credit portfolios. In: Eighth International Conference of the Society for Computational Economics. Computing in Economics and Finance, Aix-en-Provence, France
24. Sharpe W (1964) Capital assets prices: A theory of market equilibrium under conditions of risk. J Financ 19:425–442
25. Sharpe WF (1966) Mutual fund performance. J Bus 39:119–138
26. Srinivas N, Deb K (1994) Multiobjective optimization using nondominated sorting in genetic algorithms. Evol Comp 2:221–248
27. Streichert F, Ulmer H, Zell A (2003) Evolutionary algorithms and the cardinality constrained portfolio selection problem. In: Operations Research Proceedings 2003, Selected Papers of the International Conference on Operations Research (OR 2003), Springer-Verlag, pp 3–5
28. Treynor J (1965) How to rate management of investment funds. Harvard Bus Rev 43:63–75
29. Vedarajan G, Chan LC, Goldberg D (1997) Investment portfolio optimization using genetic algorithms. In: Late Breaking Papers at the 1997 Genetic Programming Conference, pp 255–263
30. Veldhuizen DAV, Lamont GB (2000) Multiobjective evolutionary algorithms: Analyzing the state-of-the-art. Evol Comp 8(2):125–147
31. Zitzler E, Künzli S (2004) Indicator-based selection in multiobjective search. In: Yao X, et al (eds) Parallel Problem Solving from Nature VIII, Springer-Verlag, Berlin Heidelberg, Lecture Notes in Comuter Science, vol 3242, pp 832–842
32. Zitzler E, Thiele L (1998) Multiobjective optimization using evolutionary algorithms - a comparative case study. In: Eiben AE, et al (eds) Parallel Problem Solving from Nature V, Springer-Verlag, Berlin Heidelberg, Lecture Notes in Computer Science, vol 1498, p 292–301
33. Zitzler E, Thiele L (1999) Multiobjective evolutionary algorithms: A comparative case study and the strength pareto approach. IEEE T Evol Comp 3(4):257–271
34. Zitzler E, Laumanns M, Thiele L (2002) SPEA2: Improving the strength pareto evolutionary algorithm. In: Giannakoglou K, et al (eds) EUROGEN 2001. Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, Athens, Greece, pp 95–100
35. Zitzler E, Thiele L, Laumanns M, Fonseca CM, Grunert da Fonseca V (2003) Performance Assessment of Multiobjective Optimizers: An Analysis and Review. IEEE T Evol Comp 7(2):117–132
36. Zitzler E, Laumanns M, Bleuler S (2004) A Tutorial on Evolutionary Multiobjective Optimization. In: Gandibleux X, et al (eds) Metaheuristics for Multiobjective Optimisation, Springer-Verlag, Lecture Notes in Economics and Mathematical Systems, vol 535