Aalto University
School of Science
Degree Programme in Operation Research

Lauri Emil Matias Nyman

A Genetic Algorithm for Generating Optimal Stock Investment Strategies

Master's Thesis
Espoo, November 23, 2017

— November 23, 2017 —

Supervisors: Professor Harri Ehtamo
Advisor: Lic.Sc. (Tech.) Aarno Lehtola

Aalto University
School of Science
Degree Programme in Operation Research

ABSTRACT OF MASTER'S THESIS

Author:	Lauri Emil Matias Nyman		
Title: A Genetic Algorithm for Generating Optimal Stock Investment Strategies			
Date:	November 23, 2017	**Pages:**	80
Major:	Systems and Operations Research	**Code:**	SCI3055
Supervisors:	Professor Harri Ehtamo		
Advisor:	Lic.Sc. (Tech.) Aarno Lehtola		

Investors including banks, insurance companies and private investors are in a constant need for new investment strategies and portfolio selection methods.

In this work we study the developed models, forecasting methods and portfolio management approaches. The information is used to create a decision-making system, or investment strategy, to form stock investment portfolios. The decision-making system is optimized using a genetic algorithm to find profitable low risk investment strategies.

The constructed system is tested by simulating its performance with a large set of real stock market and economic data. The tests reveal that the constructed system requires a large sample of stock market and economic data before it finds well performing investment strategies.

The parameters of the decision-making system converge surprisingly fast and the available computing capacity turned out to be sufficient even when a large amount of data is used in the system calibration.

The model seems to find logics that govern stock market behavior. With a sufficient large amount of data for the calibration, the decision-making model finds strategies that work with regard to profit and portfolio diversification. The recommended strategies worked also outside the sample data that was used for system parameter identification (calibration). This work was done at Unisolver Ltd.

Keywords:	Investment Strategy, Investment Portfolio, Investment Decision-Making, Genetic Optimization, Strategy Optimization, Big Data Analysis, Quantitative Investing
Language:	English

Aalto-yliopisto
Perustieteiden korkeakoulu
Tietotekniikan koulutusohjelma

DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Lauri Emil Matias Nyman		
Työn nimi: Geneettinen Algoritmi Optimaalisten Investointistrategioiden Määrittämiseen			
Päiväys:	23. Marraskuu 2017	**Sivumäärä:**	80
Pääaine:	Systeemi- ja Operaatiotutkimus	**Koodi:**	SCI3055
Valvojat:	Professori Harri Ehtamo		
Ohjaaja:	TkL Aarno Lehtola		

Investoijat kuten pankit, vakuutusyhtiöt ja yksityissijoittajat tarvitsevat jatkuvasti uusia investointistrategioita portfolioiden määrittämiseen.

Tässä työssä tutkitaan aiemmin kehitettyjä sijoitusmalleja, ennustemenetelmiä ja sijoitussalkun hallinnassa yleisesti käytettyjä lähestymistapoja. Löydettyä tietoa hyödyntäen kehitetään uusi päätöksentekomenetelmä (investointistrategia), jolla määritetään sijoitussalkun sisältö kunakin ajanhetkenä. Päätöksentekomalli optimoidaan geneettisellä algoritmilla. Tavoitteena on löytää tuottavia ja pienen riskin investointistrategioita.

Kehitetyn mallin toimintaa simuloidaan suurella määrällä todellista pörssi- ja talousaineistoa. Testausvaihe osoittaakin, että päätöksentekomallin optimoinnissa tarvitaan suuri testiaineisto toimivien strategioiden löytämiseksi.

Rakennetun mallin parametrit konvergoivat optimointivaiheessa nopeasti. Käytettävissä oleva laskentateho osoittautui riittäväksi niissäkin tilanteissa, joissa toisten menetelmien laskenta laajan aineiston takia hidastuu.

Malli vaikuttaa löytävän logiikkaa, joka ymmärtää pörssikurssien käyttäytymistä. Riittävän suurella testiaineistolla malli löytää strategioita, joilla saavutetaan hyvä tuotto ja pieni riski. Strategiat toimivat myös mallin kalibroinnissa käytetyn aineiston ulkopuolella, tuottaen hyviä sijoitussalkkuja. Työ tehtiin Unisolver Oy:ssä.

Asiasanat:	Investointistrategia, Investointiportfolio, Investointipäätösten Tekeminen, Geneettinen Optimointi, Big Data Analyysi, Kvantitatiivinen Investointi
Kieli:	Englanti

Acknowledgements

The idea and focus of this work is a result of thinking by me, Lauri Nyman and my supervisor, Aarno Lehtola. We wanted to develop a novel method to form stock portfolios. The idea got additional inspirations from existing financial sciences and converged towards the automated decision-making model of this work. We are happy that the great amount of work resulted in a model that works as desired and which can be improved and developed further.

Professor Harri Ehtamo has given me good proposals and ideas to make this thesis better. His assistance has helped much in getting this thesis into an understandable and coherent form. I would like also to notice Aarno Lehtola and my father Kai Nyman for giving their opinions about the structure of the thesis and about the good advices regarding clear textual presentation. I would like to thank these people from their help and the use of their time.

Finland, Espoo, November 23, 2017

Lauri Emil Matias Nyman

Contents

Acronyms

VAR	Value At Risk
CVAR	Conditional Value At Risk
GO	Genetic Optimization
GA	Genetic Algorithm
GP	Genetic Programming
SA	Simulated Annealing
PSO	Particle Swarm Optimization
AI	Artificial Intelligence
ROI	Return on Investment
MA	Moving Average
EMA	Exponential Moving Average
MACD	Moving Average Convergence Divergence
ROC	Rate Of Change
SMMA	Exponentially Smoothed Moving Average
RS	Relative Strength Factor
RSI	Relative Strength Index
LSQM	Least Squares Method
MDL	Minimum Description Length
MCAP	Total Market Capitalization
P/E	Price-Earning Ratio
P/S	Price-Sales Ratio
P/B	Price-Book Ratio
P/Cf	Price-Cash Flow Ratio
NYSE	New York Stock Exchange

Notations

$\mathbb{E}[\cdot]$	Expected value operator
$\mathbb{P}[\cdot]$	Probability measure
$\mathbf{1}[\cdot]$	Indicator function
$\sigma^2[\cdot]$	Variance
$\sigma[\cdot]$	Standard deviation
R_P	Portfolio return
r_{I_i}	Instrument i return
$\hat{a}$	Estimate of variable a
$\mathbf{w_I}$	Instrument allocation
$f_{penalty}$	Penalty function
U	Set of investor preferences
f	Objective function
$\mathcal{F}$	Objective functional
S	Investment strategy
$\mathcal{G}$	Individual in genetic optimization
$\mathcal{G}_{elite}$	Elite individual in genetic optimization
$\mathbf{X}_{elite}$	Set of elite individuals or strategies
$\Re(\cdot)$	Real part
$\Im(\cdot)$	Imaginary part
$\mathrm{FuzzyTrue}[a < b]$	Fuzzy less than
$\mathrm{FuzzyTrue}[a > b]$	Fuzzy greater than
$\mathrm{FuzzyTrue}[a = b]$	Fuzzy equals
$\mathrm{FuzzyTrue}[a \neq b]$	Fuzzy not equals
$\mathfrak{g}$	Model gene
S_{elite}	Elite strategy
q	Portfolio performance measure

Chapter 1

Introduction

In this work we research, design and implement a system that creates and optimizes decision-making rules for stock portfolio selection.

The implemented system creates and calibrates the strategy parameters and decision-making rules using genetic optimization. Strategies are optimized and tested using real stock market and economic data. The work was carried out at Unisolver Ltd.

The motivation for the study stems from the fact that new investment methods are highly desired. Investors including individuals, organizations, banks and insurance companies use a wide spectrum of methods to manage their investments efficiently. Large amount of resources is pointed out to develop methods for portfolio selection. In the literature the targeted balance between profit and risk vary between authors describing their methods. Better methods are constantly developed and existing methods are improved. This research constructs a new decision-making model for stock portfolio management.

1.1 Stock Investment Strategy

A stock investment strategy creation consists of selecting the candidate instruments, collecting the economic data, defining the objective function that measures the performance of a strategy or a portfolio and finally optimizing a portfolio or a strategy fulfilling the desired objective.

The selection of the optimization method is important. An optimization method must fit well into the given problem. An investment strategy is a decision-making procedure that defines the monetary amount (%) that is invested into each instrument from a given set of instruments at every moment of time. Reaching a good strategy or portfolio through optimization is a challenging task.

A wide spectrum of stock instruments can be used in portfolio selection. There are various sectors to invest in including e.g. health care, energy, industry, technology, materials, transportation and financial sectors. It is difficult to optimize strategy decision-making rules or a portfolio if there are thousands of instruments to choose from. It is better to limit the number of candidate instruments to suit into the available resources, data and knowledge.

Investment timing is crucial. The best timing is obtained when every instrument is bought at its lowest price and sold when the stock is at its peak. A good investment strategy attempts to predict these moments. Often investment strategies rely on slow moving underlying data. The stock trading decisions might consequently take place relatively seldom, e.g. weekly or monthly.

When a portfolio is formed one must be able to determine its quality. Basic portfolio quality factors are risk and return. These factors are difficult to predict but they are possible to calculate afterwards. An objective function transforms the portfolio quality factors into a single numerical figure that is used to rank different portfolios when performing portfolio selection.

There exists a wide spectrum of investment strategies. Perhaps the most trivial strategy is "buy and hold" meaning that a collection of instruments is bought and held without performing any action. The opposite strategy is a "trader strategy" where instruments are bought and sold constantly according to predefined rules. A practical investment strategy is usually somewhere between these extremes.

To make the reader familiar with stock investment portfolio selection methods this work presents a classical investment model called Markowitz portfolio model that can be used to select an investment portfolio from a given set of instruments. The Markowitz model seeks a portfolio with given amount of expected return that has the lowest possible risk.

1.2 Strategy Implementation

The system of this work uses a set of widely accepted forecasting indicators as building blocks in the decision-making rules. Recently developed portfolio selection methods are studied, and some best ideas and advantages are used in the implemented system. The strategy parameters and decision-making rules are optimized using a genetic optimization method developed for the purpose.

The emerging strategies are tested with actual historical data to get an overall picture of the developed system. The absolute performance of the formed strategies is a minor part in this work. The genetic algorithm performance is a more important factor. Still strategies having a relative good absolute performance could be achieved with the constructed structure.

1.3 Research Objectives

The first objective of this research is to (i) study models, methods and indicators that are used in investment strategies, portfolio decision-making and forecasting in financial science. The second objective (ii) is to use the best ideas and forecasting indicators to construct and document a system that forms and optimizes investment strategy decision-making rules. The third objective (iii) is to implement the documented system and to test the developed genetic optimizer performance by using real-world data.

1.4 Structure of the Work

In Chapter 2 we study features of an asset portfolio and introduce the classical Markowitz portfolio selection model. Next we discuss how to measure investment portfolio or investment strategy quality. After that some potentially efficient optimization methods for portfolio selection or model optimization are presented and compared.

Chapter 3 presents some indicators that are used for forecasting and analyzing stock instruments. Some basic principles and challenges in investment strategy and decision-making model construction are discussed.

Chapter 4 presents the implemented investment strategy generating system including the selected indicators used in the decision-making, the strategy structure, and the genetic optimization method used for the strategy optimization.

The strategies formed by the system are tested in 5 with a real test material.

Chapter 6 discusses the features of the investment strategies and the used optimization method.

The research conclusions are presented finally in Chapter 7.

Chapter 2

Investment Modeling Background

This Chapter presents the background related to the investment modeling. Section 2.1 present the basic features of an stock investment portfolio which is in this work shortened as portfolio. Section 2.2 presents the so called Markowitz portfolio model that can be used to define a portfolio fulfilling investors desires for profit and low risk. Section 2.3 discusses quantitative measures for portfolio and investment strategy quality. Next Section 2.4 presents different optimization methods that can be used while optimizing portfolios, models and stock exchange investment strategies. Finally Section 2.5 compares the performance of different optimization methods when applied to financial science problems.

2.1 Features of a Stock Portfolio

This section presents stock portfolio features including return, risk and qualitative features.

2.1.1 Return

Return measures the investment portfolio or an instrument valuation change during a certain time span (investment period). Specifically when one speaks of return it means the Return on Investment (ROI) defined for an investment period $[t, t + \delta t]$ as:

$$\mathrm{ROI} = \frac{p(t + \delta t) - p(t)}{p(t)} = \frac{p(t + \delta t)}{p(t)} - 1, \tag{2.1}$$

where $p(t)$ is the portfolio value at the beginning of the investment period and $p(t + \delta t)$ is the value at the end of the period.

ROI is often measured from the beginning of a year. Longer period results are usually expressed by average annual return.

ROI consist of two factors, the change in instrument value and the possible dividends or interest payments for the investment. The current value for a stock exchange instrument should in theory be the sum of its discounted future returns. This is why it is important to predict the company future. Current value is often less sensitive to the company's existing property, i.e. book value. Many methods are used to estimate the company's future success. Indicators used for forecasting and making portfolio allocation decisions are studied in Section 3.1.

2.1.2 Risk

Risk measures the uncertainties in the portfolio future development. It attempts to predict how probable it is that something undesired like portfolio value degradation occurs. Risk can be split into categories in several possible ways. One categorization:

- Political risk containing e.g. the effects of political decisions like changes in taxation, customs tariffs or the effects of political conflicts

- Economic risk that is mostly related to the economical performance of the target company. Economic risks include e.g. the effect of competition, product portfolio competitiveness, logistic processes, human resources and other items that can have effects to the profit and company success

- Model risk including e.g. the risk for erroneous parameter estimates, possible programming errors and the chance of model failure

Often only a few risks realize and only partially. That is why a probabilistic approach is used to measure risks. A decent numerical risk measure that fits in the situation and takes different risks into account needs to be defined. The risk analysis may differ significantly depending on the context. A nuclear power plant is designed low risks in mind[1] whereas the investment world is used to situations, where several percent of the portfolio value may be lost overnight.

The existence of various risk types makes the risk analysis challenging. In this work we take only into account the monetary risk, meaning the risk of losing capital or deviations in the portfolio value. Other risks are ignored.

Higher expected return can be achieved only by accepting higher risk, so when comparing portfolio returns, the risk must be taken into account. A portfolio with high expected profit is always inferior if there is a risk-free option like government bonds having the same profit.

The so-called risk-free-return is the return that can be achieved with no risk. In reality no risk-free investments truly exist, even governments are sometimes driven into insolvency. When speaking of risk-free-return it usually means the highest return that can be achieved with an insignificant probability for the return to vary from the expected value.

2.1.3 Qualitative Features

Investors have individual preferences for the trade-off between foreseen profit and risk. Some investor may in addition have preferences for other features like ethical factors or a desire to invest to a specific industrial sector, e.g. green energy companies. These additional factors can be taken into account, e.g, by restricting the original set of instruments to those that fulfill the necessary constraints.

2.2 The Markowitz Portfolio Model

This Section presents the widely known mathematical investment model developed by Harry Markowitz in 1952.[2] The Markowitz model presents measures for portfolio expected profit and risk. The model formulates elegantly the profit-risk trade-off for a portfolio of stock assets. When the investor's utility in risk-expected return space is known, Markowitz portfolio model determines analytically the exact instrument allocation $\mathbf{w_I} = \{w_{I_1}, ..., w_{I_N}\}$, when the set of allowed instruments $\mathbf{I} = \{I_1, ..., I_N\}$ is fixed.

2.2.1 Model Structure

The instrument's expected future return is in the original formulation assumed to coincide exactly with past returns of a given amount of time as:

$$\hat{r}_{I_i} = \mathbb{E}\left[r_{I_i}\right], \tag{2.2}$$

where r_{I_i} are the past returns and $\hat{r}_{I_i}$ is the expected future return for the instrument I_i. The sampling rate for the returns could vary from one day to up to a year. The original measure for the expected future return (2.2) assumes that the instrument's expected return remains unchanged over time.

Markowitz model risk is the standard deviation σ_P of the past returns of the portfolio. The standard deviation can be straightly derived from the variance of the past returns of the portfolio denoted as σ_P^2 and defined as:

$$\sigma_P^2 = \sigma^2 \left[r_P \right] = \mathbb{E} \left[(r_P - \bar{r}_P)(r_P - \bar{r}_P) \right]. \tag{2.3}$$

The objective is to minimize the variance (2.3) while keeping the expected return (2.2) unchanged. The problem takes a form:

$$\min_{w_{I_1},\dots w_{I_n}} \sigma_P^2 = \min_{\mathbf{w_I}} \mathbf{w_I}^{\mathbf{T}} \mathbf{\Sigma} \mathbf{w_I}, \tag{2.4}$$

so that

$$\begin{cases} \sum_{i=1}^{N} w_{I_i} = 1 & (2.5) \\ w_{I_i} \geq 0 & (2.6) \\ \sum_{i=1}^{N} w_{I_i} \hat{r}_{I_i} = \hat{R}_P. & (2.7) \end{cases}$$

In (2.4) $\mathbf{w_I}$ contains the money allocation for each instrument and $\mathbf{\Sigma}$ is the covariance matrix for the instrument returns $\{r_{I_1}, ..., r_{I_n}\}$.

The minimization of the risk must be performed so that boundary conditions presented in (2.5), (2.6) and (2.7) are satisfied. Equation (2.5) states that the weights sum up to 1, (2.6) restricts short selling and (2.7) states that the portfolio expected return $\hat{R}_P$ is fixed.

Markowitz model seeks a portfolio with a fixed expected return having the lowest possible risk measure σ_P. When the problem is solved by varying the expected return $\hat{R}_P$ one obtains a set of portfolios each having different expected returns with lowest possible risk that can be achieved with the given expected return.

The most important feature of the model is that it can lower the risk by taking advantage of the cross correlation of instrument's returns, e.g. selecting two negatively cross correlated instruments, one can lower the risk without having effect to the expected return.

The Markowitz portfolio optimization problem can be presented in an alternative form where the minimized objective function is the portfolio variance and a penalty term. One such formulation is presented by Lean Yo, Shouyang Wang and Kim Keung Lai[6] as:

$$\min_{\mathbf{w_I}} \left\{ \mathbf{w_I}^{\mathbf{T}} \mathbf{\Sigma} \mathbf{w_I} + c_{\text{penalty}} \underbrace{\left[\left(\sum_{i=1}^{N} w_{I_i} \hat{r}_{I_i} \right) - \hat{R}_P \right]^2}_{\text{penalty term}} \right\}, \tag{2.8}$$

so that

$$\begin{cases} \sum_{i=1}^{N} w_{I_i} = 1 & (2.9) \\ w_{I_i} \geq 0, & (2.10) \end{cases}$$

where c_{penalty} is the penalty factor and a quadratic penalty function is used in (2.8) to ensure that the expected return of the portfolio stays fixed. The penalty factor is ignored in the Yu, L., Wang, S. and Lai, K.K., 2008 solution ($c_{\text{penalty}} = 1$).[6]

The Markowitz portfolio optimization problem could be thought as a completely unrestricted optimization problem, if restrictions (2.9) and (2.10) are also transformed into penalty terms. In this case the objective function takes the form:

$$\min_{\mathbf{w_I}} \left\{ \mathbf{w_I}^{\mathbf{T}} \mathbf{\Sigma} \mathbf{w_I} + f_{\text{penalty}} \right\}, \tag{2.11}$$

$$f_{\text{penalty}} = c_0 \left[\left(\sum_{i=1}^{N} w_{I_i} \hat{r}_{I_i} \right) - \hat{R}_P \right]^2 + c_1 \left[\left(\sum_{i=1}^{N} w_{I_i} \right) - 1 \right]^2 - c_2 \sum_{i=1}^{N} \min \left\{ w_i, 0 \right\}, \tag{2.12}$$

where f_{penalty} is a penalty term; c_0, c_1 and c_2 are penalty coefficients.

Now we are familiar with Markowitz model formulation. Next we discuss the model preferences.

When the Markowitz portfolio optimization problem is solved by varying the expected fixed return $\hat{R}_P$, a set of portfolios is obtained denoted as $\{P_1, P_2, ..., P_N\}$ that each have the lowest possible risk that can be achieved by combining instruments (A, B, C, D) when the expected return is fixed. These portfolios form the so-called Markowitz bullet. Figure 2.1 illustrates these portfolios denoted as Markowitz portfolios.

In Figure 2.1 we have in addition convex combinations of (A, B), (B, C) and (C, D) that form portfolios. Each combination forms a curve in the risk-return graph because the instruments are correlated, for uncorrelated instruments the curve would be a straight line.

The Markowitz model does not take a stand to the risk-profit preferences of an investor. One investor may prefer more expected return and accept more risk than another. Rational investors prefer so-called Pareto optimal portfolios that settle to Pareto frontier, a curve containing all the portfolios for which the expected return cannot be increased without increasing the portfolio risk, and the risk cannot be lowered without lowering the return. Every Pareto optimal portfolio is the best possible choice of all portfolios for an investor having certain type of risk-profit preferences.

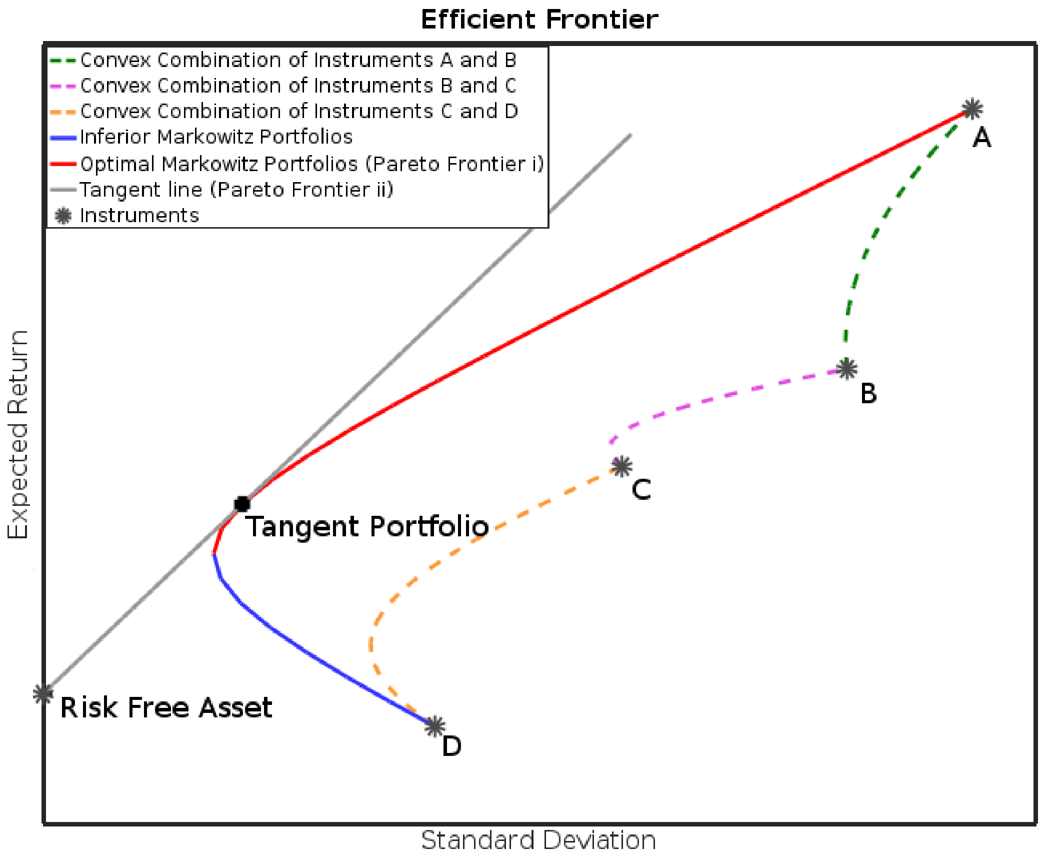

Figure 2.1: Efficient frontier for instruments A, B, C and D

It can be assumed that there are is a so-called risk-free asset having a risk-free-return. The risk free instrument is assumed to be uncorrelated with other instruments. The risk-free asset is illustrated in Figure 2.1.

Depending on the situations there are three (3) possible choices for the Pareto optimal frontier:

1. No risk-free asset: The Pareto frontier is the curve of the best possible portfolios that can be achieved by combining instruments (A, B, C, D), in Figure 2.1 the curve: Pareto Frontier (i)

2. Risk-free asset and debt available: The Pareto frontier is the curve of the best possible portfolios that can be achieved by combining instruments $(A, B, C, D, \text{Risk Free Asset})$. In Figure 2.1 the curve: Pareto Frontier (ii)

3. Risk-free asset available and debt forbidden: The Pareto frontier is the curve of the best possible portfolios that can be achieved by combining

instruments (A, B, C, D, Risk Free Asset) without taking debt. In Figure 2.1 the Pareto optimal portfolios are the portfolios setting to curves: (Risk Free Asset, Tangent Portfolio) and (Tangent Portfolio, A)

The Markowitz model is relatively simple and trivial but its concept for risk is a bit naive in the view of modern research and its method for estimating future profit is also questionable. It can create stability in long-term investing but it does not contain the necessary preferences for a strategy that targets for high profits with the lowest downside risk. Alternative risk measures are presented later in Subsection 2.2.3.

The Markowitz model estimates the expected return and variance using the classical formulas presented in (2.2) and (2.3). Variance formula (2.3) assumes that returns of a single instrument are identically distributed random variables having a normal distribution.[8] These are very strong conditions and the normality condition is shown to be false for stock returns in the study of Felipe Aparicio and Javier Estrada.[11] As a result some studies attempt to improve the Markowitz model by replacing the classical covariance and return estimates with better alternatives. In the study of Tze Leung Lai, Haipeng Xing and Zehao Chen in 2009 an alternative estimate for the covariance and return is presented.[8] It is also said that the stock market data quality is too low to apply the Markowitz model mean-variance approach.[9][10]

There exists models having more sophisticated features like self-learning and adaptation that seem to outperform the Markowitz model in both risk and return. Still the Markowitz model is a simple good classical example of an investment strategy.

2.2.2 Solution Methods

Here we present two methods, interior point method and genetic algorithm, that can be used to solve the Markowitz portfolio selection problem. The basic form of the problem is presented in Equations (2.4), (2.5), (2.6) and (2.7). The problem is a bounded quadratic optimization problem that can be solved e.g. by using interior point method also known as barrier function method. The idea is to transform the boundary conditions into barrier function, and add them to the objective function. The objective function is in this case presented in (2.13). The barrier function ensures that the boundary conditions will be fulfilled. The presentation of the method is from Aalto University course Principles of optimization.[7]

$$\min_{\mathbf{w_I}} \mathbf{w_I}^{\mathbf{T}} \boldsymbol{\Sigma} \mathbf{w_I} + \mu B(\mathbf{w_I}), \tag{2.13}$$

where B is a barrier function and μ is the barrier weight factor, suitable B could e.g. be one of the forms:

$$B(\mathbf{w_I}) := -\sum_{i=1}^{m} \ln\left[-g_i(\mathbf{w_I})\right] \text{ or } B(\mathbf{w_I}) := -\sum_{i=1}^{m} \frac{1}{g_i(\mathbf{w_I})}. \tag{2.14}$$

The interior point method requires that the boundary conditions of the original Markowitz problem presented in (2.5), (2.6) and (2.7) are transformed into form $g_i(\mathbf{w_I}) < 0$.

The interior point algorithm to solve the problem is of the form:

1. Choose an initial portfolio allocation $\mathbf{w_I^0}$ from the set of feasible portfolios. Choose $\epsilon > 0$, $\mu_0 > 0$, $\beta \in (0,1)$ and set $k = 1$.
2. Solve $\min\limits_{\mathbf{w_I}} \mathbf{w_I}^{\mathbf{T}}\mathbf{\Sigma}\mathbf{w_I} + \mu_k B(\mathbf{w_I})$. Let the solution be $\mathbf{w_I^{k+1}}$.
3. If $\mu_k B(\mathbf{w_I^{k+1}}) < \epsilon \implies$ stop.
   If $\mu_k B(\mathbf{w_I^{k+1}}) \geq \epsilon \implies$, $k \to k+1$ and go to step 2.

A method to solve the problem of step 2 is not presented. Relatively simple methods can be used to solve the step 2 problem. For a Markowitz problem a matter that makes step 2 problem easier is that Σ is positive definite in almost every real case. This is explained in the Appendix A.

A genetic algorithm is an alternative method for solving the Markowitz portfolio selection problem. The study of Hamed Soleimani, Hamid Reza Golmakani and Mohammad Hossein Salimi reveals that a decent Markowitz portfolio can be constructed for up to 2000 instruments utilizing a genetic algorithm.[19] Different optimization methods are discussed in Section 2.4. The genetic optimization is presented more specifically there in Subsection 2.4.1.

2.2.3 Alternative Risk Measures

A classical risk measure was introduced by Harry Markowitz in 1952.[2] The risk in his measure is thought to be the standard deviation of the past portfolio returns. The measure indicates how the returns have varied from the mean value during a time period. The Markowitz risk measure is defined as

$$\sigma[r_P] = \sqrt{\mathbb{E}\left[(r_P - \bar{r}_P)(r_P - \bar{r}_P)\right]} = \sqrt{\underbrace{\mathbb{E}\left[r_P^2\right] - \bar{r}_P^2}_{\text{numerically inaccurate}}}, \tag{2.15}$$

where $\sigma[r_P]$ is the standard deviation of portfolio return, r_P is the vector of past portfolio returns, $\bar{r}_P$ is the mean of past returns and $\mathbb{E}[\cdot]$ is an expected value operator. It is better to prefer the first formula of (2.15) to avoid numerical inaccuracy for a certain type of data.[3]

Markowitz risk measure (2.15) is used in various investment models to measure risk. When the only applied risk measure is (2.15) the risk caused by estimation errors for the future returns is ignored. This could have a significant impact on the risk measure functionality. Still the Markowitz risk measure is simple and suitable for many purposes.

There are alternative approaches to measuring risk. One measure is Value at risk.[4] It answers to the question about what is the minimum amount of loss in the portfolio value for a given probability p. The measure is defined as:

$$\mathrm{VaR}_p = \arg\{\mathbb{P}(X_t - X_{t+T} \geq \mathrm{VaR}_p) = p\}, \tag{2.16}$$

where p is a probability, X_t is the portfolio value in the beginning of a period $[t, t+T]$, and X_{t+T} is the value at the end of the same period; $\mathbb{P}(\cdot)$ is the probability measure. A value of $p = 5\%$ or $p = 1\%$ is often used.

An advantage of the Value at risk measure (2.16) is that a stochastic positive return does not increase the risk measure. Positive return is always desirable and not something an investor should avoid while Markowitz measure (2.15) ends up avoiding instruments that tend to react sharply to positive market news. Value at risk measure (2.16) takes roughly speaking only large drops into account. The Value at risk could in many cases be a better risk measure than a plain variance measure. The medium and small negative returns can be taken into account by using smaller values for Value at risk measure p. VaR_p can be estimated using past returns or Monte Carlo simulation.[24] A larger sample is required to sustain the accuracy of VaR_p in case of a small value of p.

Another risk measure, Conditional Value at Risk (CVaR), measures the expected loss, on condition that the portfolio has fallen by a value of VaR_p or more as:[5]

$$\mathrm{CVaR}_p = \mathbb{E}[X_{t+T} - X_t | X_t - X_{t+T} \geq \mathrm{VaR}_p], \tag{2.17}$$

in other words CVaR_p is the expected loss given that the portfolio value has fallen equal or more than VaR_p. Both CVaR_p and VaR_p are estimated similarly and their behavior has many similarities.

Low risk is preferred by actors like governments, insurance companies and banks while the profit is sometimes a less important criteria. These kind of market actors focus on diversification instead of trying to win the average market return. They analyze political and other type of risks that cannot be modeled neither with risk measures (2.15), (2.16) nor (2.17) and need to be taken into account in some other way.

2.3 Measuring Portfolio and Strategy Quality

An objective function (or functional) is used to measure numerically, with a single number, how well the investor preferences are met with a given portfolio (or strategy). If the objective function or functional is well prepared then a higher (or in some cases lower) objective value reflects a better portfolio. The objective function of the Markowitz portfolio model is presented in (2.4). The Markowitz portfolio quality is estimated using the past data, available before the selected portfolio reveals its quality.

Finding a good measure qualifying a portfolio or a strategy with a single number is a hard task because investors usually express their requirements in qualitative terms that need to be converted into a mathematical form.

Once the investor preferences are in mathematical form, it becomes possible to express the portfolio selection as an optimization problem. Subsection 2.3.1 presents how to form an optimization problem when the objective function is known. Subsection 2.3.2 presents the problem formulation when the varied objective variable is not a portfolio but an investment strategy decision-making function defining the portfolio.

The objective function is in financial science sometimes denoted as a fitness function or a utility function. In this work we denote it the objective function.

2.3.1 Objective Function

An objective function takes an instrument allocation vector (portfolio) as an argument and returns a value that measures how the investor preferences are met with the given portfolio. The objective function denoted by F is a real valued function $F(U) : \mathbf{w_I} \to \mathbb{R}$, where U contains the preferences of an investor thus defining the shape and the structure of the objective function and $\mathbf{w_I} = \{w_{I_1}, ..., w_{I_n}\}$ contains the instrument $\mathbf{I} = \{I_1, ..., I_n\}$ allocation in the portfolio.

In some problem formulations the lowest possible objective function value is targeted and others target for the highest. One can always transfer a minimization problem to a maximization problem and vice versa by performing a transformation $F \to -F$. So both the minimization and the maximization problems could be transformed to another. This section speaks for clarity reasons only about maximization of the objective value.

Finding a correct objective function depends on the case. The preferences of an investor must be defined and transformed into a mathematical form.

Optimization seeks portfolios having a high objective function value. When a portfolio has a high value for the objective function it is wise to check that the portfolio satisfies the investor needs.

The objective function is in many studies only a mathematical formula qualifying a portfolio numerically and lacks an interpretation. In that case the study may ignore the objective function value from the results and present portfolio return, risk and other preference deliverables instead. One reason for this practice is that different investors prefer widely different attributes and there are no unique objective function fitting into every investor needs.

To find a desired portfolio one must maximize the objective function value by varying the allocation $\mathbf{w_I}$ so that:

$$\mathbf{w_I^*} = \arg\left\{\max_{\mathbf{w_I}=\{w_{I_1},\dots,w_{I_n}\}} F(U, \mathbf{w_I})\right\}, \tag{2.18}$$

where $\arg\{\cdot\}$ refers to the weights $\mathbf{w_I}$ in the portfolio that are varied to maximize the objective function, $\mathbf{w_I^*} = \mathbf{w_I^*}(t)$ is a function of time covering money allocations in every moment of time. So $\mathbf{w_I^*}(t)$ covers the whole investment strategy. Note that money allocations $\mathbf{w_I}$ in (2.18) vary, while the investor preferences held in U stay constant.

Usually there are additional boundary conditions for the portfolio like the short selling restriction of the Markowitz model that must be satisfied. Common restrictions are:

1. Short selling is forbidden: $w_{I_i} \geq 0$
2. Taking debt is forbidden: $\sum_{i=1}^{n} w_{I_i} \leq 1$
3. Single asset allocation limit: $w_{min} < w_{I_i} < w_{max}$

There could be even more boundary conditions, e.g. sector allocation limit, when money allocation to specific economical sector (e.g. health care, industrial) is limited. These restrictions are reality and make the problem solving complicated. A sector allocation restriction could be modeled using binary variables, presented e.g. by K.P. Anagnostopoulos and GG. Mamanis.[12]

The additional boundary conditions could be included in the objective function. A simple approach that can be used is to give an objective value of $-\infty$ for a portfolio that violates the boundary conditions. This approach is often easy to implement but it is inefficient from the optimization point of view.

Finding the best strategy $\mathbf{w_I^*}$ from (2.18) can be a challenging task. The problem is mathematically simple, if there is only one local optima (convex problem). The objective function forms usually a structure that has a large quantity of local optima to be dealt with. Lots of optimization methods have been developed to solve these non-convex problems, some of the methods are discussed in Section 2.4.

2.3.2 Objective Functional

The approach of this work is a bit different from what is described in the Subsection 2.3.1, where a objective function was presented.

In this work an investment strategy decision-making function denoted by S, is used to select the instrument allocation $\mathbf{w_I}$, to form a portfolio. The investment strategy decision-making rules held in S is varied, affecting indirectly to the allocation $\mathbf{w_I}$. So instead of the allocation $\mathbf{w_I}$, the optimization focuses on improving the decision-making rules. This makes the objective function a functional. The objective functional is denoted by $\mathcal{F}$ and takes a form of $\mathcal{F}(U) : S \to \mathbb{R}$, where S is an investment strategy decision-making function and U are the investors preferences.

The aim is to maximize the objective functional by varying the decision-making rule S that defines the allocation $\mathbf{w_I}$, so that:

$$S^* = \arg\left\{\max_{S\in\mathcal{S}} \mathcal{F}(U)\,[S]\right\} = \arg\left\{\max_{S\in\mathcal{S}} \mathcal{F}(U)\,(\mathbf{w_I}[S])\right\}, \tag{2.19}$$

where $\mathcal{S}$ is the set of allowed rules and S^* is the rule (investment strategy) maximizing the objective functional.

In this work we use a decision-making function S to define the money allocation. The optimization concentrates on optimizing the decision-making rules. The reasons for this approach is presented later in Chapter 4 when the implemented decision-making model is represented.

2.4 Optimization Methods

In this section we present briefly genetic optimization method together with some other heuristic methods that can be used to solve different type of investment problems like Markowitz portfolio selection problem. The studied methods are in their general form focusing on their basic structure. There exist many optimization methods that combine the basic methods and enhance them. This section does not study the modified versions.

It is important to select a proper optimization method when solving a problem that has many variables and nonlinear dependencies. Many heuristic optimization methods have been developed before it became possible to process large amount of data using a computer. One optimization method may fit best to a certain type of problem, while another method outperforms it in another problem category.

2.4.1 Genetic Optimization

In the genetic optimization the so-called genome vectors $\mathcal{G}_i$ (solutions) are produced forming a basic set $\mathbf{X} = \{\mathcal{G}^1, ..., \mathcal{G}^{n+m}\}$, where every $\mathcal{G}^i$ is a candidate solution for the problem. Now one selects a given amount of best solutions from the set $\mathbf{X}$ forming an elite set $\mathbf{X}_{elite} = \{\mathcal{G}^1_{elite}, ..., \mathcal{G}^n_{elite}\}$. The solutions left in the basic set $\mathbf{X} = \{\mathcal{G}^1, ..., \mathcal{G}^m\}$ are discarded or modified. The elite solutions $\mathbf{X}_{elite}$ are combined and random modifications are made (mutations) producing new solutions that may be better than its parents, also completely new genomes may be generated.

The process is repeated to a point where no significantly better solutions are found or the solutions are refined to a desired level. When the process halts the candidate solutions for the problem are in the set $\mathbf{X}_{elite}$. There are many variations from the algorithm. In some variations no elite population exist and solutions participate in a new individual crossover with a probability proportional to the objective functional value of the corresponding individual. Often a modified version of the genetic algorithm needs to be constructed because the problem does not fit into the algorithm naturally.

Fundamental principles of the genetic optimization is presented e.g. by A.E. Eiben and M. Schoenauer in 2002.[13] Genetic optimization is often applied to problems that have a complicated form or non-continuous variables and if the optima cannot be found easily by other methods. The genetic optimization method is suitable for many purposes because it has been proved to be very robust, e.g. Jiah-Shing Chen, Jia-Li Hou, Shih-Min Wu, Ya-Wen Chang-Chien uses it in portfolio optimization.[14]

2.4.2 Simulated Annealing

The idea of the simulated annealing algorithm (abbreviated SA) originates from the physical laws of statistical thermodynamics. The algorithm is presented e.g. by Jorge Haddock and John Mittenthal in 1992.[15]

Simulated annealing algorithm starts by selecting an initial value $\mathbf{x}$, called a state. A small random change in the $\mathbf{x}$ is made, producing a value $\mathbf{x}_{\text{new}}$. The variable $\mathbf{x}$ can be either continuous or discrete. Now the transition from state $\mathbf{x}$ to state $\mathbf{x}_{\text{new}}$ occurs if for the transition probability it holds:

$$\mathbb{P}(\mathbf{x}, \mathbf{x}_{\text{new}}, T) \geq \text{random}(0, 1), \tag{2.20}$$

where T is the so called system temperature and $\text{random}(0, 1) \in [0, 1]$ is a random variable drawn from an uniform distribution.

The transition probability depends on the energy E that is in this context same as the objective function, thus $E = F$, in other words the energy for a state $\mathbf{x}$ is $F(\mathbf{x})$. A good transition probability function depends highly on application. The temperature is decreased (cooling) and the process is repeated. The system will eventually reach a thermodynamic equilibrium $\mathbf{x}_{\text{eq}}$ minimizing the energy $F(\mathbf{x}_{\text{eq}})$. Again if one wants to maximize the objective function value it can be performed by transforming F to $-F$.

Simulated annealing is a robust stochastic algorithm that usually finds good solutions. There are some drawbacks: A transition probability suiting for the application is difficult to find, and the algorithm is relative slow to converge. Finding a good solution may consequently consume a great amount of computation time.

2.4.3 Particle Swarm Optimization

Particle swarm optimization abbreviated often as PSO can be considered to imitate a swarm of birds seeking food. For an n-dimensional problem every bird i has its own global best position vector $\mathbf{p}_i = \{p_{i1}, ..., p_{in}\}$ where the bird has found the highest amount of food. The bird swarm has it's global best position vector $\mathbf{p}_g = \{p_{g1}, ..., p_{gn}\}$ where there exists the highest amount of food found by any bird. Every bird seeks food near the personal best position $\mathbf{p}_i$ but in addition every bird takes into account the globally best position $\mathbf{p}_g$.

The birds are called particles in modeling. Particle swarm optimization is introduced e.g. by Shi & Eberhart in 1998[16] and has a form:

$$v_{id}[t] = wv_{id}[t-1] + c_1 r_{i1}(p_{id} - x_{id}[t-1]) + c_2 r_{i2}(p_{gd} - x_{id}[t-1]), \quad (2.21)$$

$$x_{id}[t] = x_{id}[t-1] + v_{id}[t], \quad (2.22)$$

where $v_{id}[t]$ is the velocity of particle i in dimension d at iteration t, w is the parameter defining the trade-off between local and global search. In addition c_1 and c_2 are acceleration parameters, p_{id} is the particle i global best position in dimension d, p_{gd} is the global best position for all particles in dimension d; $r_{i1}, r_{i2} \in [0, 1]$ are randomly generated numbers.

Li & Engelbrecht in 2007 have discovered that a good convergence of particles is obtained by using values: $c_1 = c_2 = 1.49618$ and $w = 0.7298$.[17] However it is likely that good parameters are highly dependent on the problem.

Kennedy and Eberhart (1997) modified the particle swarm method to fit into problems with binary-variables[16]. This was done by replacing (2.22) with:

$$x_{id}[t] = \begin{cases} 1 & \text{if } \rho_{id} < \text{sig}(v_{id}[t]) \\ 0 & \text{Otherwise} \end{cases} \qquad \forall\ d \in \{1, ..., N\}, \tag{2.23}$$

where ρ_{id} is uniform random variable and sig$(\cdot)$ is the sigmoid function defined as sig$(x) = 1/(1 + e^{-x})$.

More information about the particle swarm optimization can be found in a work published by Hamid Reza Golmakani and Mehrshad Fazel in 2011.[18] They apply particle swarm optimization to a constrained Markowitz portfolio selection problem. Many of the conclusions above are from that study.

2.5 Comparison of Optimization Methods

In this Section we compare performance of optimization methods applied in financial science. In Subsection 2.5.1 is discussed the listed comparison criteria and Subsection 2.5.2 compares alternative optimization methods.

2.5.1 Criteria and Approach

It is necessary to apply the compared methods to same problems when doing a proper quantitative performance comparison. This is t the case in most studies in the financial science literature.

It seems that publications measure the performance of a certain set of optimization methods by utilizing a too narrow set of test material, in addition the structure of the optimized object varies between publications, e.g. the optimized object can consist of a Markowitz portfolio selection method, fuzzy logic or decision trees or other type of structures. A certain optimization method is efficient in optimizing a specific kind of model or decision-making rules. This is why the quantitative performance differences do not show up. That is a reason why we discuss the performance qualitatively.

Optimization methods seems to find a solution in a relatively short time. Various studies speak about seconds or minutes. E.g. study of Hamed Soleimani, Hamid Reza Golmakani and Mohammad Hossein Salimi reveals that a decent Markowitz portfolio can be constructed for up to 2000 instruments utilizing genetic algorithm in 17 minutes. For 500 instruments the optimization converges in 6 to 8 minutes.[19] So the solving time is not a significant factor because stock transaction decisions are often made in a time scale of days. A time span of hours might be sufficient even for fast trading. Solving time can become a restrictive factor only if instrument selection needs to be made among a set of thousands of instruments.

The major focus is usually to obtain an investment portfolio that is as good as possible, while computing power is a secondary issue that is solved by adding the necessary processing capacity. This turned out to be the case also when the optimization method for this work was selected.

2.5.2 Comparison in the Literature

Hamid Reza Golmankani and Mehrshad Fazel (2011) extend the Markowitz portfolio selection problem and solve it using a modified particle swarm optimization and compare it to genetic optimization with a test sample varying from 9 up to 150 stocks.[18] The return r is fixed and the portfolio variance σ^2 varies. The study found that in this case the modified particle swarm optimization outperforms the genetic optimization. The genetic optimization seems to find good solutions when the portfolio is small. The modified particle swarm optimization is superior when the portfolio is large. The superiority was demonstrated with the capability to find portfolios with better (min σ^2), mean variance $E[\sigma^2]$ and deviations in the portfolio variances ($\mathrm{var}[\sigma^2]$). The mean solving time was also shorter. One must note that the modified particle swarm optimization method was tailored for this purpose, increasing the available solving time might have cut the differences between the compared methods.

Tsung-Jung Tsai, Chang-Biau Yang and Young-Hsing Peng (2010) have implemented an investment strategy based on the so called *technical indicators* explained later in Section 3.1. The indicators include e.g. moving average (SMA), global trend indicator (GTI) and monitoring indicator (MI).[20] The method uses trigger signals to make buying and selling decisions. The method includes parameters (weights) that are set and adjusted to obtain calibrated decision-making rules. The performance of the method is compared with static user-defined weights, static genetic optimized weights and dynamic weights optimized with a genetic algorithm. The comparison is carried out with a large test sample, a set of instruments varying from 35 to 45 having a time series history of roughly 9 years. The dynamically optimized method gave the best (ROI) performance followed by the genetic optimized static method. The user-defined method was minor.

Ying-Hua Chang and Ming-Sheng Lee have developed in 2015 a trading strategy based on Markov Chains.[21] Market moves are split into separate categories based on the days return. The return can e.g. have 15 categories between $(-7\%, 7\%)$ each denoted as $n \in (1, 15)$. The model assumes that the category of the previous day return predicts the category for the next day return. In the model return of category i is assumed to be followed by category j with a probability of p_{ij}. The method parameters are optimized separately using simple and modified genetic algorithm. The study compares different trading strategies based on average return of 24 periods settled evenly in 2003-2014. The modified genetic algorithm outperforms in this case the simple genetic algorithm. The optimization performed by simple genetic algorithm achieves marginally higher average return than a Buy & Hold strategy for TWN50 (TWSE Taiwan 50 Index) and TAIEX (Taiwan Capitalization Weighted Stock Index). A noticeable thing is that a model optimized with modified genetic optimization can achieve positive returns even when model optimized with the two other methods have negative returns. The method is based on Markov chains that has many similarities with forecasting with autocorrelation.

Somayeh Mousavi, Akbar Esfahanipour and Mohammaed Hossein Fazel Zarandi (2013) constructs a trading system that is based on decision trees.[22] A unique decision tree is constructed for each stock I_i in the set of stocks $\mathbf{I}$ to take into account each stock's unique individual behavior. So the method consists of decision trees D_i for each instrument I_i. Each decision tree is optimized independently using genetic algorithm to obtain efficient trading rules. Each decision tree D_i has adjusted allocation w_i as an output for stock I_i. The allocations w_i are finally normalized so that the method takes partly into account the instruments interdependences. The method is tested and compared to a 2-year performance of a Buy & Hold strategy for (TEPIX) Tehran Stock Exchange and S&P500 during 2009-2011. When only returns are considered, the decision tree outperforms both Buy & Hold, TEPIX and S&P500 indexes significantly with a level of significance varying between 1 % -5 %. In addition the decision tree model portfolios achieve a better risk measure, and not only the relative but also the absolute performance of the decision tree model seems to be excellent.

2.5.3 Optimization Summary

New optimization methods have been developed and existing methods have been improved, e.g. there are many variations from the genetic optimization method. Stochasticity can be added to optimization to make a method more robust. Modifications made to optimization algorithms can show advantages in certain types of problems and have disadvantages in others, e.g. convex problems can usually be solved efficiently with simple methods but are solved inefficiently with stochastic methods like genetic optimization or simulated annealing. Stochastic methods often outperform in non-convex problems because deterministic methods are easily stuck in local optima. The problem type is consequently valuable information and one should pay attention to it when selecting an optimization method, e.g. covariance matrix of the Markowitz portfolio selection problem is in reality always positive definite (Appendix A). This is valuable information while selecting an optimization method for solving the Markowitz portfolio selection problem.

New optimization methods are often developed by applying and modifying an optimization method that appears to work in the real nature. Imitating natural phenomena, like in genetic algorithm and simulated annealing, seems to work well providing robust and efficient methods that fit to many applications.

2.6 Conclusions

We studied in this chapter the portfolio features including risk and return. An example of an analytical investment model called Markowitz portfolio selection model familiarized the reader with investment modeling. We presented the objective function and functional that are used to provide a numerical measure of the investor utility. Simple optimization methods used in financial science were presented and finally the investment methods and their optimization found in the literature were discussed. At this point the reader has become familiar with mathematical modeling and optimization in the financial science.

Chapter 3

Domain Modeling

In this chapter the reader is familiarized with some of the common indicators and methods for financial forecasting. The basic principles and decision-making indicators used in investment strategies are introduced and the practical challenges related to missing data are discussed.

3.1 Indicators for Forecasting

In this section we present some of the most common *fundamental indicators* used to describe the company economic situation and some *technical indicators* that are derived from the stock price behavior. The most appropriate indicators are selected to the implementation presented in Chapter 4.

The selected subset of the indicators having supposed predictive value for stock assets future return (ROI), and risk (volatility, drawdown) should be as independent from each other as possible. Using a subset of indicators that depend on each other is not appropriate. Unnecessary and harmful redundancy, called collinearity, is caused if one indicator can be constructed using the others.

It seems that authors of financial science have developed similar or even linearly dependent indicators. That is why selecting a subset of independent indicators is a major challenge and must be performed with care. The set of forecasting indicators should be sufficiently large to ensure a decent prediction but at the same time as limited as possible. The used method or decision-making model always finds a good fit if there are too many indicators from which the optimization method can pick a combination. The method would seem to adapt fine but would rarely work outside the sample data used to calibrate the parameters.

3.1.1 Economic Indicators

Some common *fundamental indicators* including MCAP, P/E, P/S, P/B and P/Cf that are used to forecast behavior and to indicate fundamental factors of a company. Explanations for these indicators:

- MCAP is the company total valuation in currency
- P/E is the price-earnings ratio. The ratio of market value and annual earnings of the company
- P/S is the price-sales ratio indicating how large annual sales does the company have with respect to its market value
- P/B is the price-to-book ratio indicating the ratio between the market value of the company and its balance sheet value (book value)
- P/Cf is the ratio of market value and annual cash flow

A low price compared e.g. to the book value (P/B) or revenue (P/S) is a hint towards the possibility that the company might be in income or cash flow related difficulties.

The values of the relative indicators (P/S, P/Cf, P/E and P/B) fall into a certain range that depends on the sector, type of the company and interest level. It is possible e.g. that a large industrial company has plenty of capital invested to machinery and investors might believe that the only valuable thing in the company are the machines. As result the stock valuation P of the firm is near its book value B the price-book ratio being $P/B \approx 1-2$. A start-up company could be an opposite example, the invested capital might be low but the valuation of the company is high as a result of high expectations regarding the business idea. A start-up firm could consequently have a very high P/B-ratio.

3.1.2 Technical Indicators

Technical indicators that can be derived from the stock market data are typically based on stock price charts or stock turnover histograms. The definition of many *technical indicators* originate from the era when calculations had to be performed by hand explaining the simple definitions. More sophisticated indicators have shown up later but not all of them provide advantage over the simpler ones. This subsection focuses on the most commonly used basic *technical indicators* alongside with few relatively new and potentially useful ones.

Stock value moving average MA(n) also known as simple moving average SMA(n) is an indicator that defines the average price of the stock over last n days. The moving average indicator smooths out rapid changes in the stock price. MA can be thought to present the slowly moving component of the stock price. It is defined for a stock valuation time series p as:

$$\mathrm{MA}\left[p_t\right](n) = \frac{p_t + p_{t-1} + ... + p_{t-n+1}}{n} = \frac{1}{n}\sum_{i=0}^{n-1} p_{t-i}, \tag{3.1}$$

where p_t is the day t price, and n is the time window for the moving average in days.

The exponential moving average (EMA) is a widely used alternative for MA. EMA is a moving average with exponentially decreasing weights for the past prices. EMA gives the trend of the stock in pretty much similarly as MA, but responds more quickly to rapid changes. It is defined as:

$$\begin{aligned}\mathrm{EMA}\left[p_t\right](\alpha) &= \frac{p_t + (1-\alpha)p_{t-1} + (1-\alpha)^2 p_{t-2} + ...}{1 + (1-\alpha) + (1-\alpha)^2 + ...} \\ &= \frac{\sum_{i=0}^{\infty}(1-\alpha)^i p_{t-i}}{\sum_{i=0}^{\infty}(1-\alpha)^i} \approx \frac{\sum_{i=0}^{n}(1-\alpha)^i p_{t-i}}{\sum_{i=0}^{n}(1-\alpha)^i},\end{aligned} \tag{3.2}$$

where $\alpha \in (0,1)$ is a coefficient defining how much the past effects. If α is near 1 then the past is ignored and the indicator adapts immediately to changes. If α is near 0 then past has a large effect and the indicator is approximately the average stock value from a longer period of time.

Moving average convergence-divergence (MACD) measures the so-called stock momentum and has similar properties as the classical derivative operator. Calculating MACD for a stock valuation time series p is done by generating a time series denoted by p' that is the difference between two EMAs with EMA coefficients α_1 and α_2. The MACD is the EMA of p' with an EMA coefficient of α_3. Using the formula in the appendix B, MACD can be formed as:

$$
\begin{aligned}
\mathrm{MACD}[p_t](\alpha_1,\alpha_2,\alpha_3) &= \mathrm{EMA}\left[\mathrm{EMA}\left[p_t\right](\alpha_1) - \mathrm{EMA}\left[p_t\right](\alpha_2)\right](\alpha_3) \\
&= (1-\beta_3)\sum_{n=0}^{\infty}\left[\left(\frac{1-\beta_1}{1-\dfrac{\beta_3}{\beta_1}}\right)\beta_1^{n+1}\right] - \left(\frac{1-\beta_2}{1-\dfrac{\beta_3}{\beta_2}}\right)\beta_2^{n+1} \\
&\quad -(1-\beta_3)\sum_{n=0}^{\infty}\left[\left(\frac{1-\beta_1}{1-\dfrac{\beta_3}{\beta_1}} - \frac{1-\beta_2}{1-\dfrac{\beta_3}{\beta_2}}\right)\beta_3^{n+1}\right] p_{t-n},
\end{aligned}
\tag{3.3}
$$

where $\alpha_1 = 1-\beta_1$, $\alpha_2 = 1-\beta_1$ are the EMA-coefficients for the original series p_t and $\alpha_3 = 1-\beta_3$ is the EMA-coefficient for the p'_t series.

From (3.3) one can note that MACD is effectively a weighted difference between 3 EMAs of the original time series. It indicates the stock direction (stock momentum).

Another momentum indicator called rate of change (ROC) is defined for a stock with price p_t as:

$$
\begin{aligned}
\mathrm{ROC}\left[p_t\right](n) &= 100\cdot\left(\frac{p_t}{p_{t-n}} - 1\right) \\
= 100\cdot\frac{p_t - p_{t-n}}{p_{t-n}} &= 100n\cdot\frac{\mathrm{MA}[p_t](n) - \mathrm{MA}[p_t](n-1)}{p_{t-n}}.
\end{aligned}
\tag{3.4}
$$

ROC is relatively similar to MACD. Using both indicators simultaneously could result in unnecessary complexity.

The Discrete Fourier transformation (DFT) or the computationally efficient version of it called the Fast Fourier transformation (FFT) is used to divide a time series into different sine and cosine waves. In exact terms it is not an indicator but it can be used to analyze stock behavior and reveal phenomena. DFT can be best understood by presenting a time series x_t using sine and cosine function as:

$$
x_t = \sum_{k=0}^{N-1}\left[\Re(X_k)\cos\left(2\pi i t\frac{k}{N}\right) + \Im(X_k)\sin\left(2\pi i t\frac{k}{N}\right)\right], \tag{3.5}
$$

where $\Re(X_k)$ is the real, and $\Im(X_k)$ is the imaginary part of the discrete frequency spectrum $X = \Re(X) + i\Im(X)$ at k. The spectrum X reflects the frequencies present in the original time series.

There is one fundamental problem when DFT is applied in financial science. The stock market is closed during the weekends and during some other holidays. These missing days change the phase of the original time series x leading to a misleading Fourier analysis. The Fourier analysis can for this reason be reliably used only for low frequency detection. Lower frequencies are less sensitive to short term discrepancies.

Another commonly used indicator called Relative Strength Factor (RS) is defined as:

$$\mathrm{RS} = \frac{\mathrm{SMMA}\left[U_t\right](n)}{\mathrm{SMMA}\left[D_t\right](n)}, \tag{3.6}$$

where $U_t = \max\{\mathrm{close}_t - \mathrm{close}_{t-1}, 0\}$, $D_t = \min\{\mathrm{close}_{t-1} - \mathrm{close}_t, 0\}$, and SMMA are exponentially smoothed moving averages with smoothing factor $\alpha = 1/n$. The exponential smoothing SMMA is defined for x_t as $s_t = \alpha x_t + (1-\alpha)s_{t-1}$, where $s_0 = x_0$. The relative strength factor (RS) is often converted to a so called relative strength index (RSI) defined as:

$$\mathrm{RSI} = 100 - \frac{100}{1+\mathrm{RS}},\ \mathrm{RSI} \in (0, 100). \tag{3.7}$$

According to many investors RSI indicates weather the stock is oversold (low value) or overbought (high value). The RSI includes nonlinear operations and that is one reason why it might contain new information not included in other indicators.

Current literature introduces a large spectrum of different methods for portfolio selection. The arguments and indicators used in the studies are surprisingly similar. There are various other indicators including for example volume oscillator and momentum oscillator. These indicators are ignored in this work.

Ayca Cakmak Pehlivanli's, Barik Asikgil's and Guzhan Gulay's research paper in 2016 augments machine learning to select a suitable set of indicators that forecasts the next day returns for stocks in the Istanbul Stock Exchange (ISEX). Appropriate forecasting indicators are selected from a comparatively large set of 97 indicators. The most promising indicators for prediction were price momentum, relative strength index, stochastic oscillator, the demand index, intraday momentum index, the random walk index, ultimate oscillator, William R and the commodity channel index. These indicators were selected by minimizing the next day return forecasting error for the instruments. Their approach differs from the more commonly used forecasting period of one or several months. It is possible that the Istanbul stock exchange behavior differs from large trading hubs like New York Stock Exchange (NYSE), but the research reveals anyhow potentially very valuable information about candidate forecasting indicators.

3.2 Principles in Investment Strategy Construction

A good method is as simple as possible but still fulfill its purpose. Degrees of freedom in the method should be small compared to the sample size used for the parameter identification phase. The reason for this is that when adjusting e.g. decision-making rule parameters the system will always fit into the sample if there are too many degrees of freedom. The system would seem to perform fine but would rarely work outside the parameter identification sample data.

Attempts have been made and methods have been developed to include the simplicity as a target into the automated method building. One interesting approach for the problem is developed by Jorma Rissanen in 1978.[23] His method is called minimum description length principle (MDL). The MDL seeks a method or a model that stores parameters and estimation errors in the smallest possible memory. When the total information is minimized the method is automatically as simple as possible still fulfilling its purpose. The implementation of the MDL principle could however be relatively complex depending on the situation; implementation is ignored in this study but it could be a part of the future development plan.

3.3 Practical Challenges

This section presents the practical challenges related to model construction.

3.3.1 Instrument Set Selection

Selecting a decent set of instruments is a major task when optimizing an investment strategy. Even an excellent model can suggest bad portfolios if it has only inferior instruments to choose from. Sufficient time series data for the stock prices and for *fundamental indicators* need to be available for testing. A time series may have some incompleteness that needs adjusting before an analysis can be performed, missing data points need to be filled up.

Stock price history may have forecasting value. If additional indicators e.g. trading volumes and fundamental data is needed, they need to be collected. It may be necessary to discard some potentially good candidate instruments due to poor availability of the necessary data.

3.3.2 Collecting the Required Indicators

Fundamental indicators need often to be collected manually from the interim reports. Manual work is however tedious when speaking of hundreds of instruments with a long history. Some missing data points might be obtained by interpolation but it may be better to completely ignore an indicator if the data series is too incomplete.

3.3.3 Filling the Missing Information

Interpolation is a commonly used method for filling the missing data points of an incomplete time series. Imagine that a time series x has data points x_t and x_{t+2}, but the point x_{t+1} is missing. The missing point can be interpolated using the existing data. Linear interpolation is the most common choice. It has in this case the form $x_{t+1} = (x_t + x_{t+2})/2$. Nonlinear interpolation methods can be used in cases when the time series is nonlinear. The choice of an interpolation method has only minor effect to the result if the interval is small between the points and a change in x is small. To avoid unnecessary complexity linear interpolation is used in this work. One must notice that interpolation does not generate new information and it is only done to fill the missing data to ensure the method operation.

Interpolation works if the concerned phenomena is continuous. A so-called stock split sometimes breaks the continuity of stock price time series. In a stock split the value of a stock is adjusted to a convenient level if the stock price has risen by a significant amount. In a stock split the owner may e.g. have his 10 shares converted to 100 shares with a 1/10 of the original value, investment total value does not change but the value of a stock abruptly decreases to 1/10 of its original value. Un-continuous points need to be located and fixed before analysis is performed.

3.3.4 Company Dividends

Some funds and companies do not deliver profit through dividends, they may e.g. buy their own stocks instead. The stock price development includes in these cases the dividends, that are automatically taken into account.

Most of the companies pay dividends as a result of a successful year or as an old habit. Dividends may take place once a year or several times during the year as is common on the United States markets. Dividends are usually a significant factor for the portfolio return and are recommended to be included in a model. Including dividends into a model is a time-consuming task that often requires manual work. Simple methods may ignore the dividends and the results are consequently less perfect.

In our decision-making model we have used a time series data that is split adjusted but the data is not dividend adjusted, thus dividends are ignored. This decision made things simpler and reduced the manual workload significantly. This work focuses on creating and analyzing the optimization, while the decision-making model performance is a secondary priority. Adding dividends into the decision-making model could be part of the future development.

3.3.5 Transaction Costs

The transaction costs are always present when buying and selling stock shares. These costs may include the contract with a stock broker, the cost per transaction and the cost per exchanged share. The costs can be simulated using predefined transaction costs.

The investment world deals usually with transactions greater or equal than \$ 10000. With these lots and if the trading frequency is moderate the transaction costs are relatively small and can be ignored. This work resulted in strategies where the trading is performed once a month and the total amount of money in the simulation is \$ 100000 so the transaction costs are relatively small and could be ignored.

3.3.6 Dealing with the Outliers

The data could include outliers caused by erratic readings or by other reasons. Instruments with far outlying data should be corrected or removed if possible.

Deletion or correction is not always possible. If the outliers are not caused by errors and if it is necessary to include the outliers in the analysis then it is important to ensure that the outliers do not dominate the results. A single erratic data point can for example distort a least square fit (LSQM) significantly. Some methods are less sensitive to outliers. The median is for example a less sensitive indicator (robust indicator).

There were relatively few outliers but relative many systematic problems in the raw data used in this work and it was important to ensure that the outliers were dealt properly.

3.4 Conclusions

This Chapter studied the indicators used for forecasting stock behavior and for portfolio selection decision-making. Some principles that should be kept in mind when constructing an investment method were discussed. Problems related to imperfect and missing data were studied. The reader has now the essential knowledge regarding investment model building for practical use.

Chapter 4

Strategy Representation and Optimization

This chapter introduces the constructed and implemented generic structure used to generate decision-making rules to form stock portfolios. A single individual decision-making rule is called an investment strategy.

In Section 4.2 we present an overview of the structure. Section 4.3 discusses types of objective functionals that can be used to score the decision-making rules. The implemented genetic optimization method used to calibrate the rules is presented in Section 4.4.

The structure and the decision-making system and the genetic optimizer is implemented and integrated into Unisolver Ltd's *Unisolver Portfolio Risk Manager*. To make the simulations and optimization possible the heavy calculations are carried out using custom software functions built for this purpose with *C++ programming language*. The decision-making rule generation and optimization is implemented using *LISP symbolic programming*.

The strategies (decision-making rules) are scored with an objective functional. The final goal is not to find portfolios but a set of strategies producing portfolios with a good performance. This is a conceptual difference to the Markowitz portfolio model approach. Using genetic algorithm results in an automated system for finding decision-making rules that define an investment portfolio. These rules can be stored as accumulated knowledge about what has worked in the past.

4.1 Indicators for Decision-Making

Moving averages (SMA and EMA) of stock prices, relative strength index (RSI), price standard deviation $\sigma[p]$ and basic *fundamental indicators* are selected to be used in the decision-making as forecasting indicators. This set of indicators is selected so that redundancy is avoided as much as possible while simultaneously maintaining a high level of versatility.

The used indicators are not plain linear predictive factors. They are basic building blocks for the decision-making rule that selects a set of stocks to the portfolio. The decision-making rule gets its formula during optimization and it may result in having a multitude of dependencies including e.g. linearity, bi-linearity and exponential behavior.

All the used *technical indicators* have at least one argument, e.g. SMA and EMA have an integer number of days as argument. The arguments are randomly generated from an exponential distribution and rounded to a nearest integer. Exponential distribution has user defined suitable intensity parameter for every used indicator. The exponential distribution is chosen because it is natural to calculate e.g. moving averages for 5, 7, 11, 20, 50, 100 and 200 days. These numbers fit well into an exponential distribution. Geometric distribution is an alternative option for the exponential distribution.

4.2 Structure of the Investment Strategy

This section presents the generic structure of the decision-making rule starting from single factors and progressing to the overall structure of the rule. Figure 4.1 illustrates the overall procedure of the decision-making.

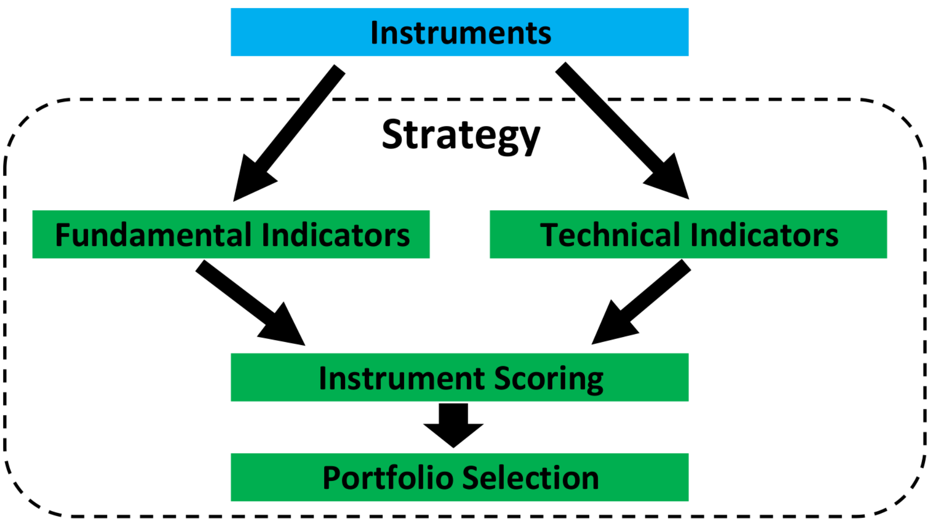

Figure 4.1: Investment Strategy Structure

4.2.1 Fuzzy Operators

The decision-making relies strongly on indicator comparison. Indicators are compared to other indicators or to pure numbers using a specific kind of comparison operators, denoted in this work as fuzzy operators. The operators are fuzzy equal, fuzzy unequal, fuzzy less than and fuzzy greater than. They have forms:

$$\text{FuzzyTrue}[a = b] = \frac{\mathbf{1}\left[ab > 0\right] - \mathbf{1}\left[ab < 0\right]}{1 + \left[-\epsilon_{=} \frac{2|a| - 2|b|}{\max\left\{|a| + |b|, \delta_{=}\right\}}\right]^2}, \tag{4.1}$$

$$\text{FuzzyTrue}[a \neq b] = \text{sign}\left(\text{FuzzyTrue}[a = b]\right)\left(1 - |\text{FuzzyTrue}[a = b]|\right) \tag{4.2}$$

$$\text{FuzzyTrue}[a < b] = \frac{1}{1 + \exp\left[-\epsilon_{<}\left(\frac{2a - 2b}{\max\left\{|a + b|, \delta_{<}\right\}}\right)\right]}, \tag{4.3}$$

$$\text{FuzzyTrue}[a > b] = \frac{1}{1 + \exp\left[-\epsilon_{<}\left(\frac{2b - 2a}{\max\left\{|a + b|, \delta_{<}\right\}}\right)\right]}, \tag{4.4}$$

where $\epsilon_{<}, \epsilon_{=} > 0$ are the sensitivity parameters, $\delta_{<}, \delta_{=}$ are parameters needed to ensure that when negative values are compared to positive values, or if $|a + b|$ is a very small number, the divider differs enough from zero. We can define the functions by using limits $\delta_{<}, \delta_{=} \to 0$.

A set of fundamental boundary conditions are used in the fuzzy operator construction. The conditions are presented in the following equations:

$$\text{FuzzyTrue}[a < b] = \text{FuzzyTrue}[b > a], \tag{4.5}$$

$$\text{FuzzyTrue}[a < b] + \text{FuzzyTrue}[a > b] = 1, \tag{4.6}$$

$$\text{FuzzyTrue}[a = b] + \text{FuzzyTrue}[a \neq b] = 1 \ \ \forall\ ab > 0. \tag{4.7}$$

$$\text{FuzzyTrue}[a = b] + \text{FuzzyTrue}[a \neq b] = -1 \ \ \forall\ ab < 0. \tag{4.8}$$

$$\text{FuzzyTrue}[a = b] + \text{FuzzyTrue}[a \neq b] = 0 \ \ \forall\ ab = 0. \tag{4.9}$$

Equation (4.5) ensures that arguments ordering does not have effect to the comparison result. In the special case when $a = b$ both fuzzy operators (4.3) and (4.4) take a value of 1/2. Equation (4.6) ensures that the sum of fuzzy-variables indicating a being less and greater than b is 1. Equation (4.7) ensures that a equal b and unequal b sums to a fuzzy-variables being 1.

It is not possible to say if a pure number is large or small. It depends on the situation and context. That is why the comparison functions are in addition scale-invariant satisfying relations:

$$\text{FuzzyTrue}[a < b] = \text{FuzzyTrue}[Ca < Cb], \tag{4.10}$$

$$\text{FuzzyTrue}[a = b] = \text{FuzzyTrue}[Ca = Cb], \tag{4.11}$$

$$\text{FuzzyTrue}[a \neq b] = \text{FuzzyTrue}[Ca \neq Cb], \tag{4.12}$$

Conditions (4.10), (4.11) and (4.12) are met if a and b are not too near to zero or to each other. This can be seen from the structure of functions (4.1), (4.2), (4.3) and (4.4). This is a problem that cannot be avoided in a relative comparison. However all of the relations hold for all (a, b) if $a, b \neq 0$ and $a \neq b$ when $\delta_<, \delta_= \to 0$.

It can be said e.g. that numbers 10 and 11 are relatively near to each other, as opposite when comparing e.g. 0 and 1 or -2 and 2 it is unclear how near the numbers are in relation to each other. Only the absolute difference can be presented. When constructing a relative magnitude comparison function these problems must be dealt with. The method of this work partially deals with this problem by producing negative fuzzy values when comparing negative numbers to positive ones, also value zero is always fuzzy unequal to every nonzero number. This can be mathematically expressed as:

$$\text{FuzzyTrue}[a = -a] = -\text{FuzzyTrue}[a = -a] \ \forall \ a, \tag{4.13}$$

$$\text{FuzzyTrue}[0 = b] = 0 \ \ \forall \ b, \tag{4.14}$$

$$\text{FuzzyTrue}[a \neq -a] = -\text{FuzzyTrue}[a \neq -a] \ \forall \ a \neq 0, \tag{4.15}$$

$$\text{FuzzyTrue}[0 \neq b] = 0 \ \ \forall \ b. \tag{4.16}$$

The presented few compromises were necessary in the fuzzy operators, while making them suitable for practical applications. Still the functions are as formal as possible.

Figures 4.2, 4.3, 4.4 and 4.5 present the fuzzy comparison operators. It can be seen that the fuzzy operators are pairwise symmetric. The problem in the definition is the singular point where $a, b \to 0$ and the relative difference have no exact definition. Lowering the $\epsilon_<$ and $\epsilon_=$ values makes the functions less steep and increasing makes the operator steeper. A good parameter choice depends perhaps on the application. The testing revealed that $\epsilon_<, \epsilon_= \in [4, 10]$ is a good choice at least for our strategy structure and test material.

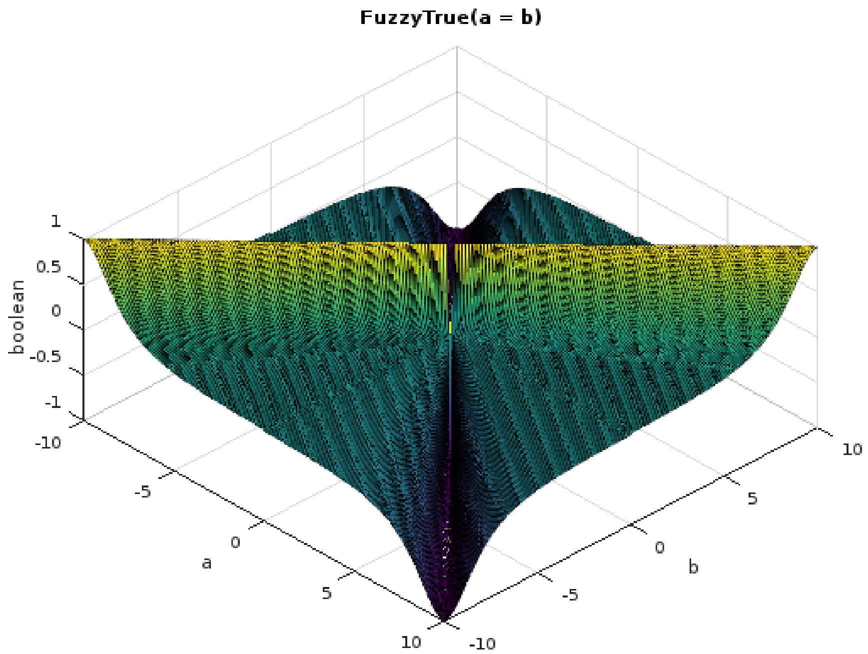

Figure 4.2: Fuzzy functions having parameters $\epsilon_< = \epsilon_= = 5, \delta_< = \delta_= = 0.01$.

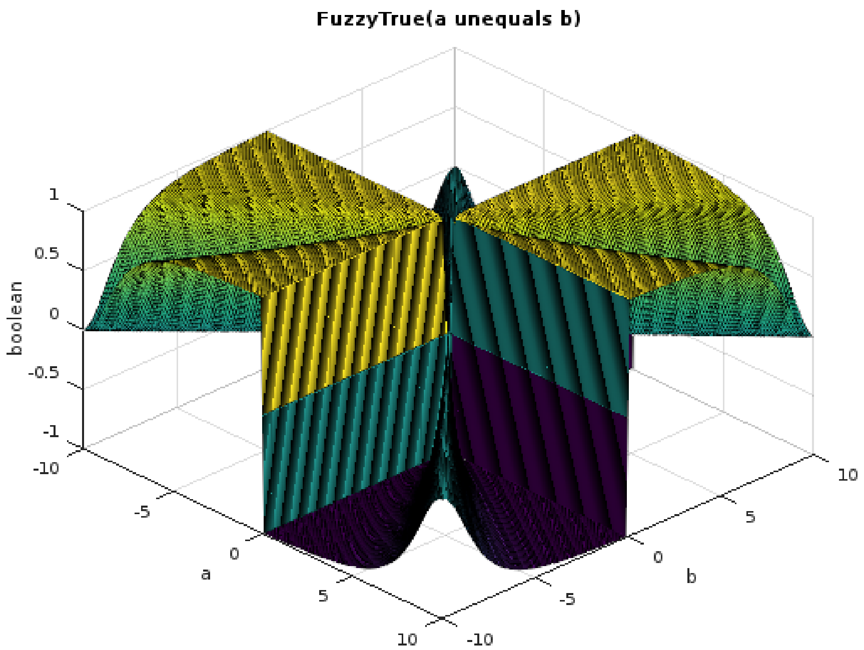

Figure 4.3: Fuzzy functions having parameters $\epsilon_< = \epsilon_= = 5, \delta_< = \delta_= = 0.01$.

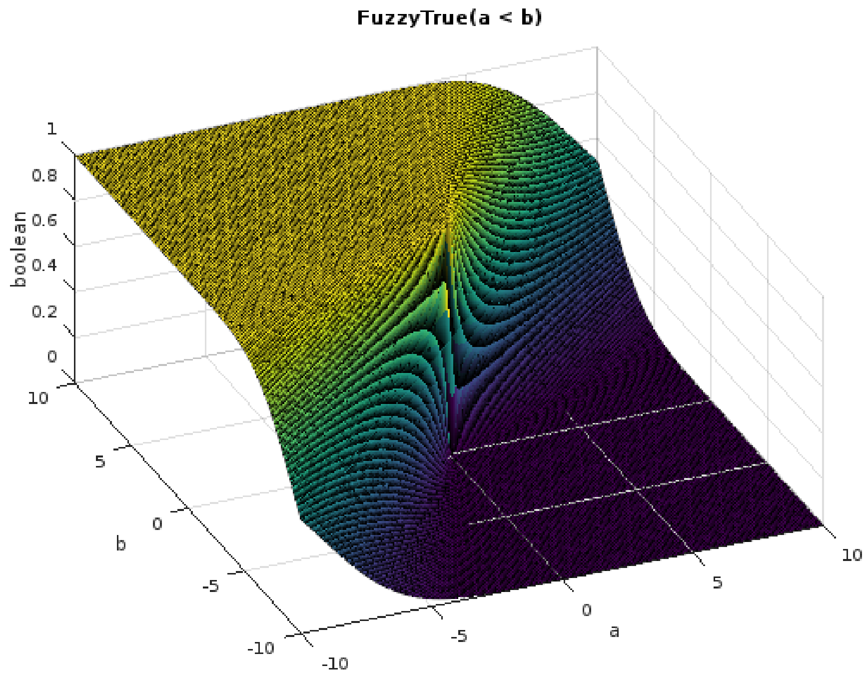

Figure 4.4: Fuzzy functions having parameters $\epsilon_< = \epsilon_= = 5, \delta_< = \delta_= = 0.01$.

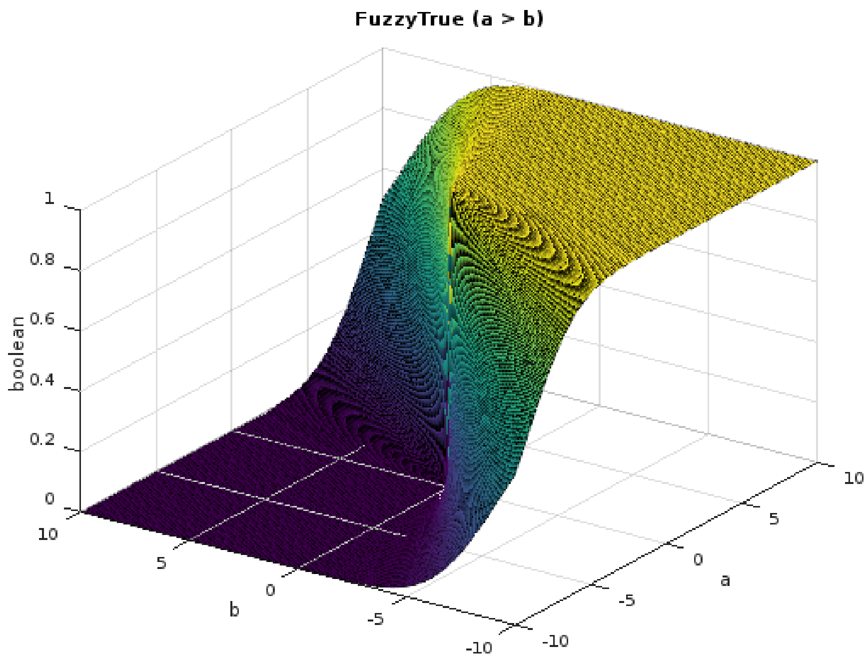

Figure 4.5: Fuzzy functions having parameters $\epsilon_< = \epsilon_= = 5, \delta_< = \delta_= = 0.01$.

4.2.2 Strategy Gene

The basic element of the decision-making rule is called a gene. A gene performs a fuzzy comparison for a pair of certain predefined type of indicators. The comparison is of the form of (4.1), (4.2), (4.3) or (4.4). A decision-making rule contains several different types of genes each comparing different type of attributes. A gene of type i is denoted by $\mathfrak{g}^i$. A gene is actually a function that has a fuzzy number as an output, e.g. genes of type l and m denoted as $\mathfrak{g}^l$ and $\mathfrak{g}^m$ can be of the form of:

$$\mathfrak{g}^l = \text{FuzzyTrue}[\text{RSI}(10) > 0.5], \tag{4.17}$$

$$\mathfrak{g}^m = \text{FuzzyTrue}[\text{SMA}(15) < \text{Price}]. \tag{4.18}$$

The output of a single gene is always a fuzzy variable having a value between $[-1, 1]$. So the value can be also negative due to the structure of fuzzy equals (4.1) and unequals (4.2) comparison operators.

4.2.3 The Overall Picture of the Strategy

A set of genes presented in Subsection 4.2.2 are collected together to form the decision-making rule to select the instruments. In this work a single rule defining the instruments is called a strategy and is denoted by S. There are exactly N_{genes} number of genes in each strategy. The strategy type is defined by its number depending on how many similar type of genes are used. A strategy S has the form:

$$S = f\left[\mathfrak{g}_1^{\text{Type}(1)}, ..., \mathfrak{g}_{N_{genes}}^{\text{Type}(N_{genes})}\right], \tag{4.19}$$

where f defines how the instrument scoring and weighting is calculated from a group of genes, so f defines the form of the decision-making rule. It could be e.g. simply an additive function of the basic elements $\mathfrak{g}_i^j$.

The function f can be a simple minimum (fuzzy and), a maximum (fuzzy or), a sum or product of the fuzzy elements $\mathfrak{g}_i^j$ or some custom-made function. If a traditional fuzzy function (and/or) are used then the negative fuzzy values produced by the genes $\mathfrak{g}_i^j$ must be dealt with. This can be done for example by ignoring the use of fuzzy equal (4.1) and unequal (4.2) functions in the strategy. In a demonstrating example (4.20) f is a weighted sum function and there are exactly one type of each gene thus in that case Type$(i) = i$ and the strategy, i.e the decision-making rule to select the instruments has a form:

$$S = \sum_{i=1}^{N_{genes}} c_i \mathfrak{g}_i^{\text{Type}(i)} = c_1\mathfrak{g}_1^1 + c_2\mathfrak{g}_2^2 + c_3\mathfrak{g}_3^3 + \cdots + c_{N_{genes}}\mathfrak{g}_{N_{genes}}^{N_{genes}}, \tag{4.20}$$

where c_i presents the weighting of each gene. If the model is still simplified so that the weighting is ignored ($c_i = 1\ \forall i$) a strategy reduces to a form that can be e.g. of the form:

$$S = \sum_{i=1}^{N_{genes}} \mathfrak{g}_i^i = \text{FuzzyTrue}[\text{SMA}(15) < \text{P}] + \cdots + \text{FuzzyTrue}[\text{MCAP} > \$1B]. \tag{4.21}$$

The strategies generated in this work contain exactly one gene of each type. This approach prevents such redundancy where similar indicators participate several times in the instrument scoring. This work uses a scoring function f of the same generic form for every strategy. Still the scoring function contains parameters that can be different for each strategy. The set of allowed parameters is limited to avoid too many degrees of freedom in the strategy.

Equation (4.21) maps a single instrument to a number. The numbers, called instrument score, are calculated separately for each instrument. The higher the instrument score is, the better it is in the sense of strategy S. A strategy selects a predefined number of best scored instruments forming an investment portfolio. The instruments could be equally weighted, or the weighting could depend on the instrument scoring. In this work the strategy selects a set of 20 instruments which it considers to be the best. A decent choice in general is to select 10-25 best scored instrument and to invest equally in each.

Figure 4.6 illustrates a set of k strategies each having the same number of genes. Each strategy covers one individual investment strategy, a decision-making rule that defines the instruments.

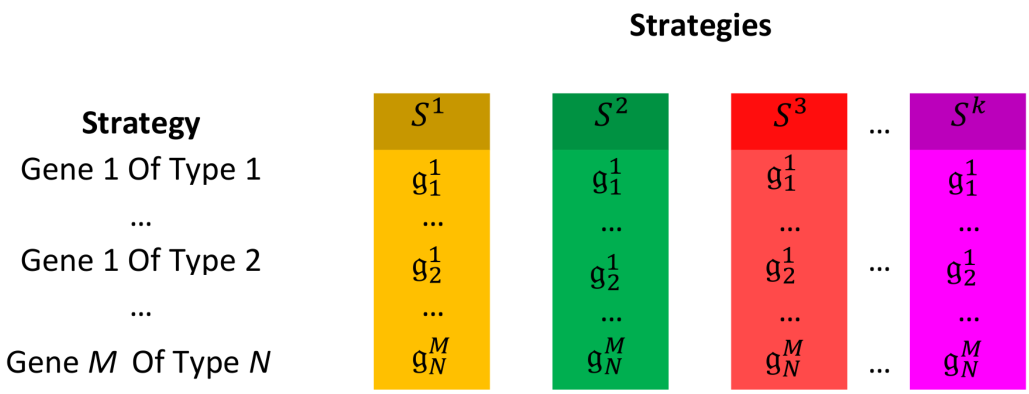

Figure 4.6: A set of strategies

Strategies update their portfolio on monthly basis. A strategy S can be expressed as:

1. Calculate the scoring for each instrument using strategy scoring function $f\left[\mathfrak{g}_1^{\mathrm{Type}(1)}, ..., \mathfrak{g}_{N_{genes}}^{\mathrm{Type}(N_{genes})}\right]$

2. Select e.g. the 10 best scored instruments with $w_i \approx 10$ % to form a portfolio

3. When a month have passed, repeat steps 1 and 2; Note that the available capital is different in each step

The procedure presented above forms a different investment portfolio for each month that can be e.g. the following:

20070928	(adc.us 159)	(alg.us 203)	(cbd.us 329)	$\cdots$	(usph.us 337)
20071031	(adc.us 150)	(alg.us 224)	(cbd.us 302)	$\cdots$	(usph.us 331)
...	...	...	...	...	...
20111031	(adc.us 412)	(agm.us 488)	(cbd.us 421)	$\cdots$	(yum.us 183)

Table 4.1: Portfolios

In the Table 4.1 the rows presents the portfolios, each of which are bought at the start of the month presented in the first column. Instruments are presented as pairs that have a ticker first and the number of shares as second.

4.3 Strategy Objective Functional

An objective functional $\mathcal{F}$ also called fitness function in genetic optimization, attempts to measure how well the aspects of the investor preferences are met with the given strategy S.

Strategies presented in Section 4.2.3 select the instruments using the past data. The strategy quality on the other hand is defined from the portfolio behavior that is selected: the quality is defined from the portfolio behavior after it is bought. This is one significant difference between our strategy and the Markowitz portfolio method that defines the quality from the portfolio past return and risk.

The objective functional measuring the strategy quality consists in this work of the strategy portfolio annualized return, Sharpe ratio, maximum drawdown and a few other measures. Possible mathematical forms for the objective functional are discussed next.

Different aspects and strategy attributes can be taken into account by using a weighed sum of the portfolio factors. In this case objective functional is of the form:

$$\mathcal{F}(U)\,[S] = \sum_{i=1}^{n} c_i q_i \left[\mathbf{w}_{\mathbf{I}}(S)\right], \tag{4.22}$$

where c_i is a weight for portfolio performance measure q_i, $\mathbf{w}_{\mathbf{I}}$ is the allocation of instruments. A performance measure has a positive weight when it is favorable ($c_i > 0$), and a negative weight ($c_i < 0$) when it is undesired. The weights must be properly calibrated to prevent one performance measure from dominating.

An alternative form for the objective functional is:

$$\mathcal{F}(U)\,[S] = \prod_{i=1}^{n} \left\{q_i \left[\mathbf{w}_{\mathbf{I}}(S)\right]\right\}^{p_i}, \tag{4.23}$$

where $p_i \in \mathbb{R}$ is the power factor defining the weighting of different features. If $q_i < 0$, fraction powers must be set so that no complex numbers are produced. When a feature q_i is favorable, $p_i > 0$, and if it is not favorable, then $p_i < 0$.

According to our experience, when the power factors n_i are adjusted correctly the objective functional of the form (4.23) seems to become less sensitive for the preferences q_i than (4.22). Strategies optimized using an objective functional of the form (4.23) seems to prefer performance measures more evenly and achieve a better qualitative performance than the strategies optimized using objective functional of the form (4.22).

A bit risk averse approach was used while the objective functional was constructed. In other words, a strategy producing a good return, relatively high Sharpe ratio and having no significant drawdown was preferred.

One can note that the decision-making system ignores the interdependencies between the selected instruments during the instrument selection stage. For example the cross correlation is neglected from the structure in (4.19). This is one disadvantage in our strategy compared to e.g. Markowitz model that efficiently uses the instruments cross-correlation in the decision-making. Still a well working objective functional can take the cross-features implicitly into account, e.g. Sharpe ratio improves if the variance of the portfolio returns is low. Targeting a better Sharpe ratio results in portfolios where the cross correlation is implicitly taken into account, leading automatically to diversified portfolios.

4.4 Implemented Genetic Optimization Method

In this section we introduce the implemented genetic optimization method used to optimize the strategies presented in Section 4.2.3. Subsection 4.4.1 of this section illustrates the overall structure of the optimization method, Subsection 4.4.2 presents the operations from which the optimization method consists of and Subsection 4.4.3 presents a pseudo-code covering the whole optimization process.

4.4.1 Genetic Optimization Structure

Figure 4.7 illustrates the overall structure of the investment strategy optimization.

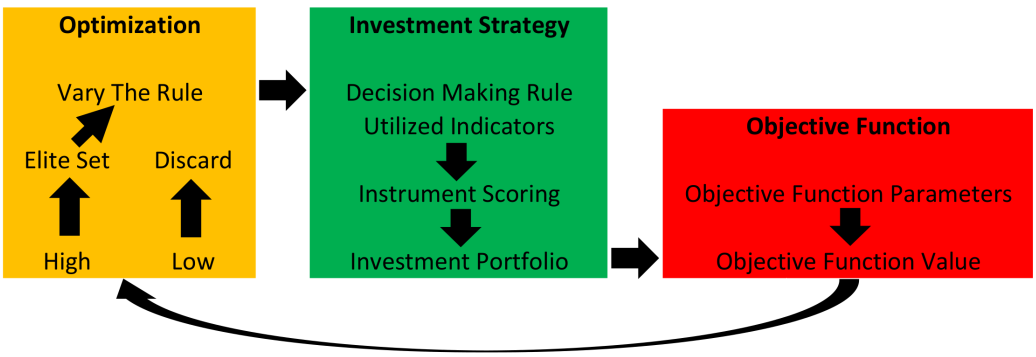

Figure 4.7: Overcall picture of applying genetic optimization

First a set of investment strategies is generated, and each strategy selects portfolios as presented in the Table 4.1. The objective functional value for each strategy is calculated based on the portfolio behavior. Strategies having high objective function values are stored, and improved in the genetic optimizer.

4.4.2 Description of the Genetic Operations

The goal is to use genetic optimization to find well working strategies and discard the poor ones. This subsection presents the set of basic operations that the genetic optimization algorithm uses.

An elite group of strategies is denoted by $\mathbf{X}_{elite} = \left\{ S^1_{elite}, ..., S^{N_{elite}}_{elite} \right\}$ where the best found individuals are held. There are three kind of operations used during the optimization: Generation of new strategies (A), crossover of strategies (B) and mutation of strategies (C):

(A) Generation of a new strategy: A new pure random strategy having N_{genes} random genes is generated

(B) Crossing a new strategy: 2 random parents S^i_{elite} and S^j_{elite} from the group of elite strategies $\mathbf{X}_{elite}$ are selected and a new strategy is generated by mixing the parent genes. The operation leaves $\mathbf{X}_{elite}$ unchanged

(C) Mutation: One random strategy S^k_{elite} from group $\mathbf{X}_{elite}$ is selected and $n_{mutation}$ random genes are replaced with new random ones, generating a new mutated individual S^k. The operation leaves the original parent S^k_{elite} unchanged in the set $\mathbf{X}_{elite}$

Figure 4.8 illustrates the elite set, and a new set that is composed from mutated, crossed and completely new strategies. The numbers in the genes refer to the gene type and the colors distinguish genes of different individuals. Note that the mutated and crossed individuals have parents in the original elite group, marked with colors.

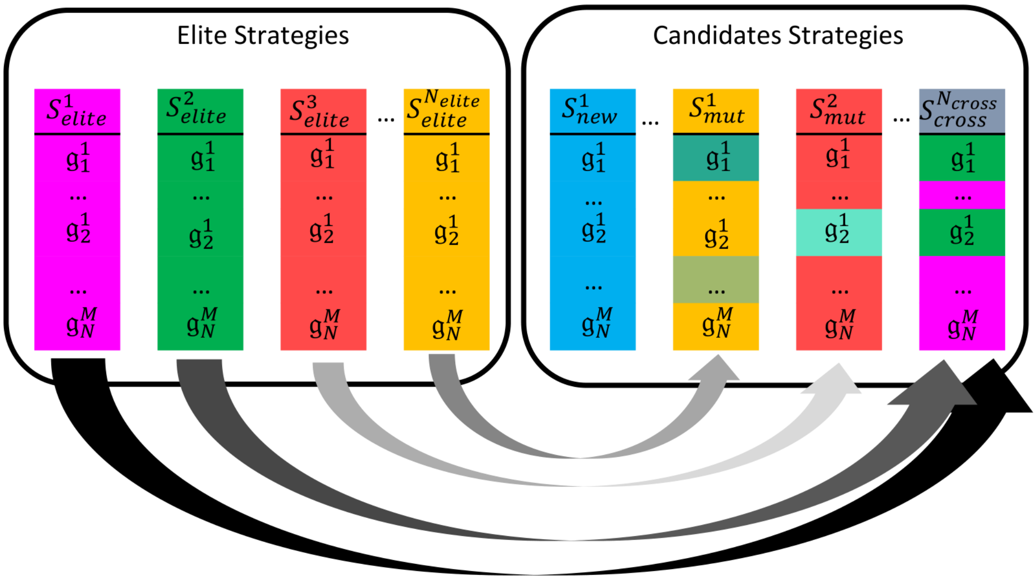

Figure 4.8: Genetic Algorithm

4.4.3 Genetic Optimization Pseudo-Code

The optimization pseudo-code is presented below. The optimization begins with population initialization and is followed by the population optimization.

1. N_{elite} number of random strategies are generated forming the elite set $\mathbf{X}_{elite} = \left\{S^1_{elite}, ..., S^{N_{elite}}_{elite}\right\}$

2. (a) N_{cross} times random parents S^i_{elite} and S^j_{elite} from $\mathbf{X}_{elite}$ are selected and a new individual is crossed from the parents

   (b) $N_{mutated}$ times a random strategy S^i_{elite} is selected from the set $\mathbf{X}_{elite}$ and $n_{mutation}$ random genes are replaced with new genes producing a new strategy

   (c) N_{new} completely new random strategies are generated

3. Strategies generated in the stage 2 are merged to the elite group $\mathbf{X}_{elite}$

4. $\mathbf{X}_{elite}$ is cut so that N_{elite} number of strategies having the highest objective function value are left to the set forming a new elite set. Other strategies are discarded

5. Steps 2, 3 and 4 are performed total $N_{iterations}$ times

When the optimization stage is complete the set $\mathbf{X}_{elite} = \left\{S^1_{elite}, ..., S^{N_{elite}}_{elite}\right\}$ that is obtained as an output from the optimization loop form the final strategies. Optimization should be continued until the strategies in the set $\mathbf{X}_{elite}$ fulfill the investor's desires (objective functional value is high enough).

4.5 Conclusions

We presented in this chapter the implemented system that forms the decision-making rules (investment strategies) for selecting stock investment portfolios. Two possible forms for the objective functional were discussed and the genetic optimization method used to optimize the strategies was presented.

Chapter 5

Strategy Testing

In this chapter we test and evaluate the implemented decision-making system presented in Section 4.2. Strategies are optimized and tested with a predefined test material and the portfolios the strategies selects are presented and discussed. The strategies are optimized using the genetic algorithm according to Section 4.4. In the optimization stage an objective functional having a formula of the form presented in Section 4.3 is used to score the strategies.

Section 5.1 introduces the used test material. After the test material is introduced we begin to present the simulation results. In the first test the number of iterations applied for the strategy optimization is varied and the effect (number of iterations) are discussed in Section 5.2 by presenting the optimized strategies portfolios. After that the optimization method convergence is discussed in Section 5.3. Section 5.4 tests the strategy optimizer when the parameter identification and validation periods is varied. Sensitivity of the strategies with respect to decision-making parameters are tested in Section 5.6. In the final test presented in Section (5.7) the strategies are modified by separating the *fundamental analysis* and *technical analysis* factors into two separate decision-making models. Both of these models are optimized separately and the performance of the strategies is discussed.

The constructed decision-making model can be compared to other portfolio selection models. This can be done by comparing the portfolios of the formed strategies to portfolios formed by e.g. Markowitz method. The Markowitz strategy fits into long-term investing and the strategies in this work are mainly meant for short-term investing and to adapt into monthly changes. Markowitz model objective function is different and its value is defined from the past behavior of the portfolio, while the objective function value for strategies in this work is defined from the realized features of the portfolio. Conclusion is that the portfolios of the strategies in this work and Markowitz portfolios have different objectives and they are not fully comparable. Our work uses a significant quantity of instruments requiring major amount of implementation and computing time. Building a Markowitz Portfolio using the same set of instruments would be a large task because the of huge amount of data, and it would require more suitable algorithm to be developed. The numerical comparison of these methods has for this reason been omitted.

Many different settings and test setups can be surveyed. The parameter identification stage, objective function and strategy parameters can be varied, and different amount of optimization can be applied. There are limitless options for testing but only a relatively limited amount of test results is presented.

5.1 Data Used in the Simulations

This section presents the test material, and the use of it in the strategy optimization and simulation.

5.1.1 Test Material

As a test material we use the daily stock data of New York Stock Exchange (NYSE) from June 2009 to March 2017 covering roughly a time span of eight years. The set of candidate stock instruments consists of more than 2000 individual companies. The daily stock data is collected from one source only. Using only one data source is an advantage, if the data contains systematic errors, it is likely that the genetic algorithm forms strategies that adapt to these errors.

It was recognized that the set contains companies in a liquidation state causing unrealistic disturbances and irrational outliers. There were also companies having zero daily trading volumes meaning that those particular instruments cannot be exchanged in reality. These problems were solved by limiting the set of instruments to those having a market cap (MCAP) value of higher than \$ $500M$. A company having a market cap of \$ $500M$ is still relatively small in the NYSE stock exchange context. After the limiting we still had a good sample of 1378 companies left for the simulation.

5.1.2 Training and Validation Periods

The time series data from June 2009 to March 2017 was split into two periods. The first period is used in the strategy optimization (parameter identification phase). The first period spans from June 2009 to November 2012. Approximately first 100 days of the first period could not be used in the strategy simulation because the data from those days was needed while calculating the indicators used for the instrument selection, e.g. evaluating MA(100) requires 100 last stock values from the history. The strategies having the highest objective functional value was selected from the parameter identification period (the first period), resulting in the elite set of strategies.

The second period spanning from November 2012 to March 2017 was used to validate if the strategies are working outside the sample. Outside of the sample testing is important to ensure that the strategies are not by change fit into sample.

5.2 Effect of Refining

Figures 5.1, 5.2, 5.3 and 5.4 illustrate the portfolio progress of 10 strategies having the highest objective functional values at the identification period, while in each figure a different amount of optimization is applied. Both the parameter identification and the validation periods are presented separately. To understand the portfolio performance, S&P 500 index consisting of the 500 largest North American companies is added as a reference series. The investment portfolio value is calculated monthly and also the exchange decisions (buying and selling) of the stock shares are carried out monthly, similarly as we presented in Table 4.1.

5.2.1 Unoptimized Strategies

Figure 5.1 illustrates the return on investments of 10 randomly generated unoptimized strategies. The upper figure illustrates the strategies during the parameter identification period and the lower one illustrates the strategies during the validation period.

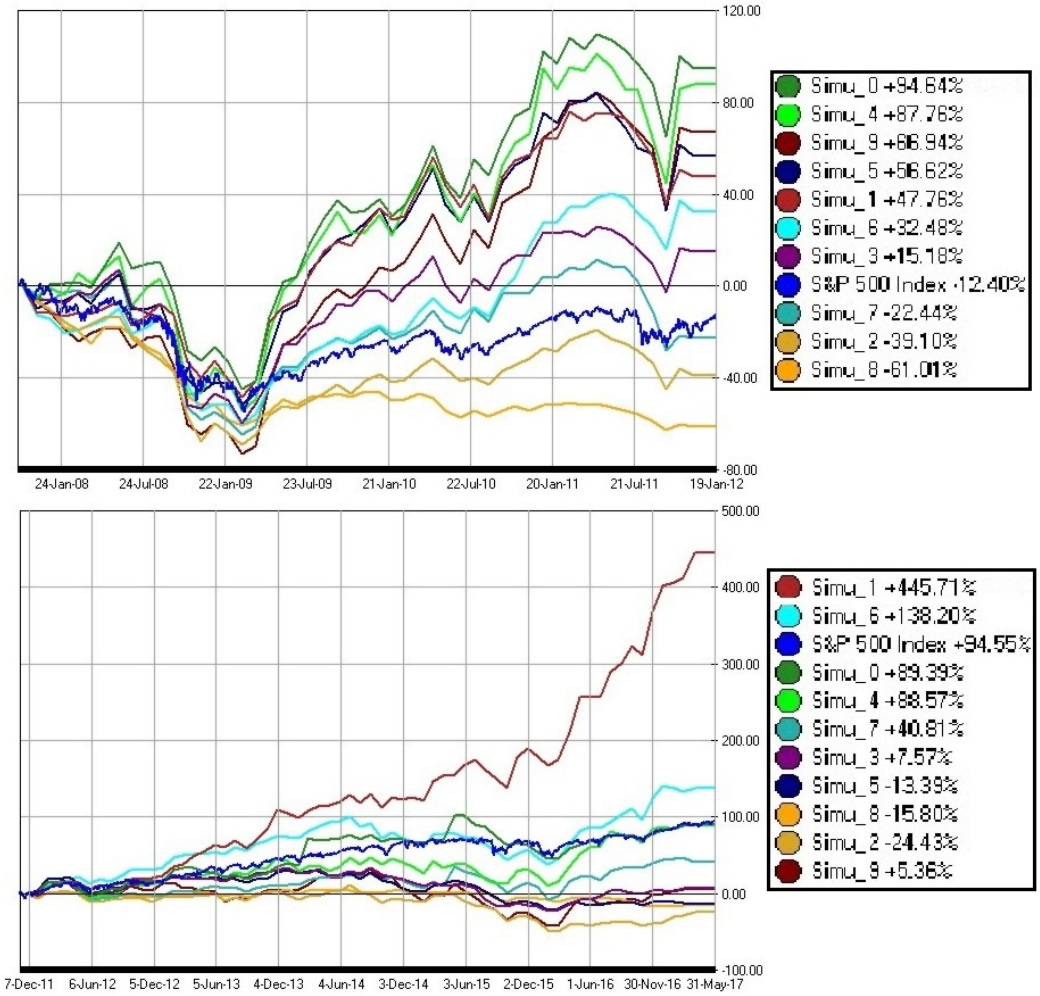

Figure 5.1: Return on investment (%), strategies unoptimized

The randomly generated strategies produce very unstable portfolios compared to the S&P 500 Index. There is a huge deviation between the strategy portfolios. One must note that these portfolios are selected by randomly generated unoptimized strategies.

5.2.2 Marginally Optimized Strategies

The first chart in Figure 5.2 illustrates the strategy portfolios during the parameter identification period and the next chart illustrates the same strategies during the validation period. The strategies are marginally optimized. The total of 135 strategies has been evaluated during the optimization stage.

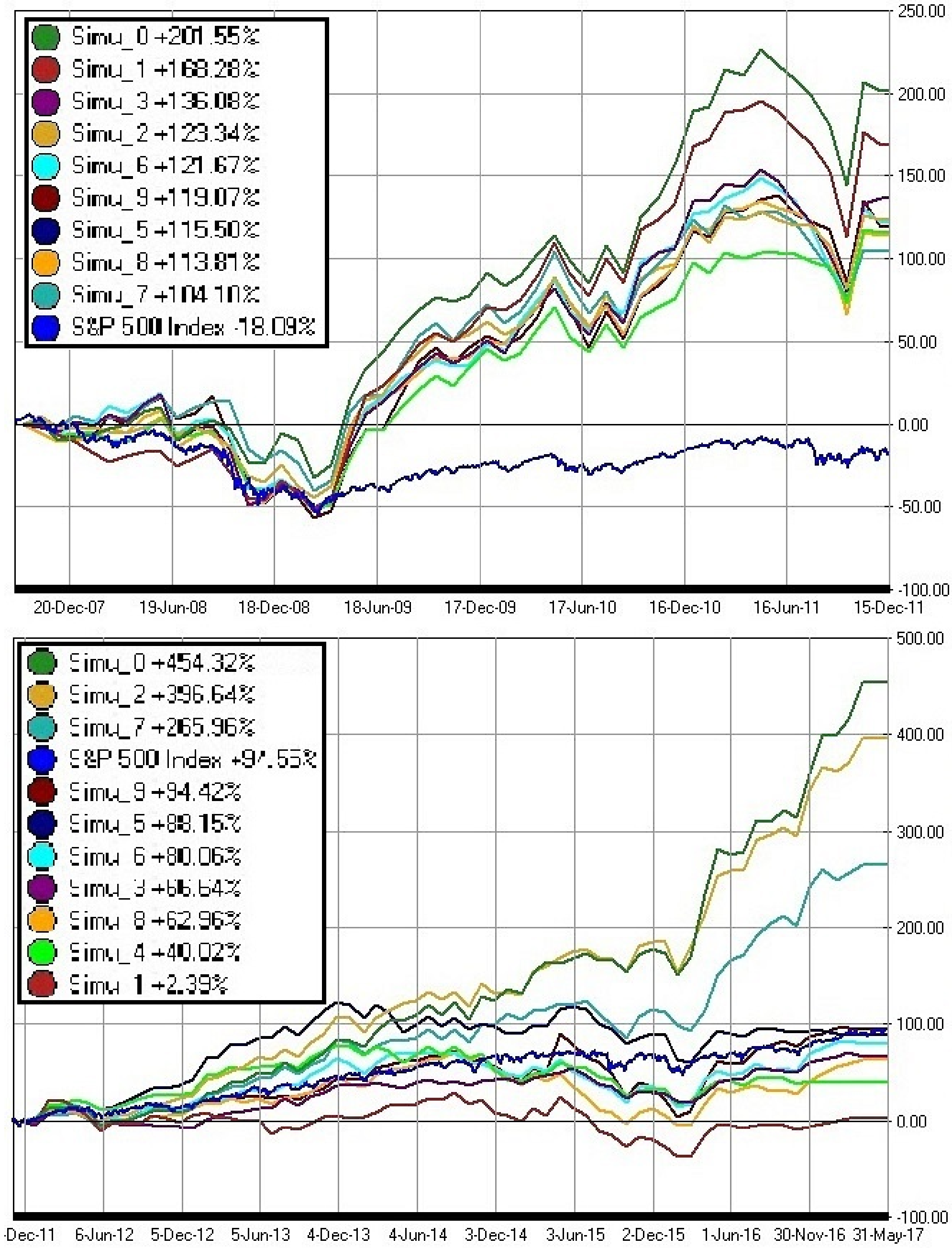

Figure 5.2: Return on investment (%), strategies marginally optimized

The strategies possess better performance than the unoptimized strategies presented in Figure 5.1. Every strategy outperforms the S&P 500 index during the parameter identification period. There are significant deviations in the strategies performance during the validation period.

5.2.3 Medium Optimized Strategies

At this stage, 675 strategies have been evaluated during the optimization.

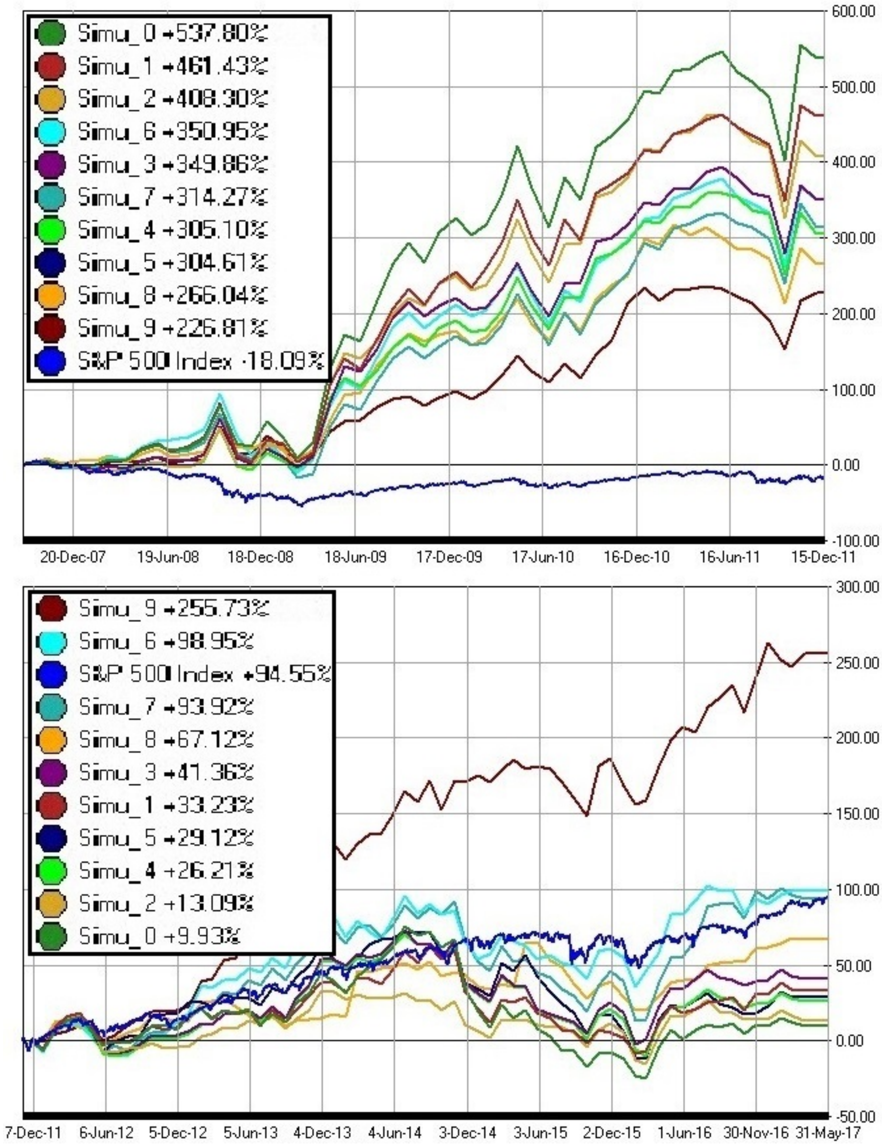

Figure 5.3: Return on investment (%), strategies medium optimized

Figure 5.3 is similar to the figure presented before. The S&P 500 Index has been outperformed significantly during the parameter identification period but the generated strategies seem to achieve low performance during the validation period. This indicates that the strategies are a fitting into the sample and do not perform outside the sample.

5.2.4 Highly Optimized Strategies

Figure 5.4 illustrates the return on investments of 10 well optimized strategies. In the process 2550 strategies have been evaluated and 10 best strategies have been selected.

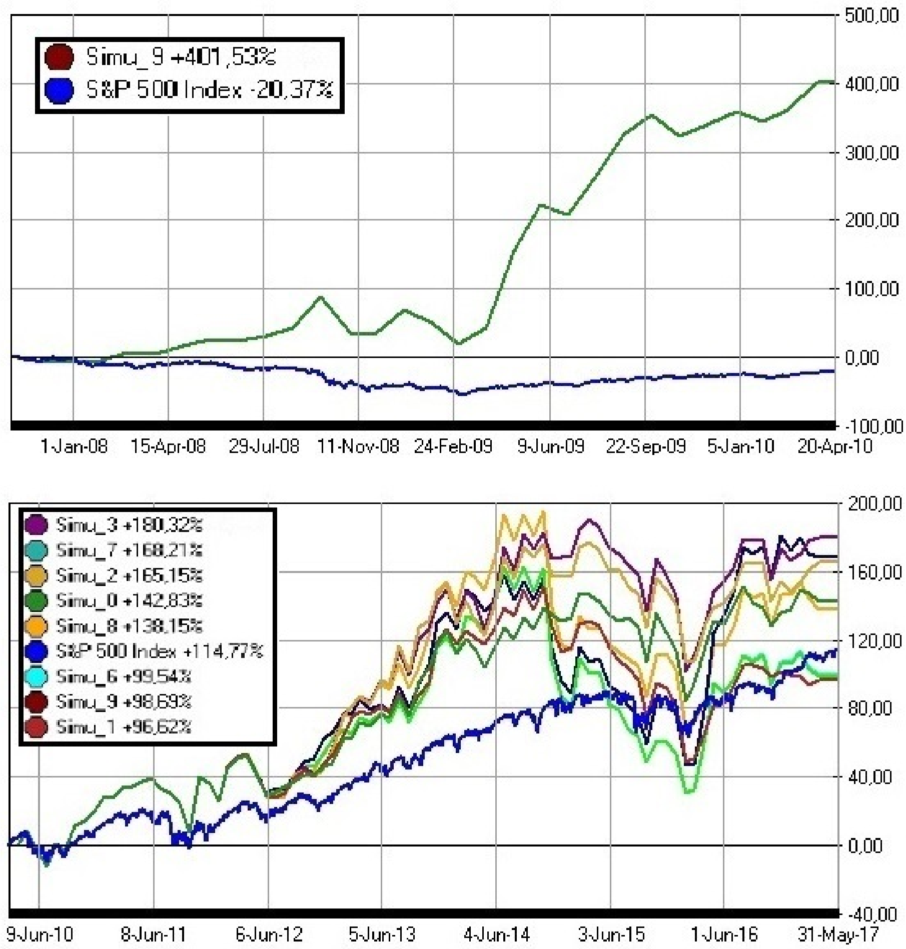

Figure 5.4: Return on investment (%), strategies highly optimized

One can note that in Figure 5.4 the best 10 elite strategies selected the same portfolio during the identifying period and their performances are consequently identical. The overwhelmingly good performance during the identification phase indicates a fit into the sample. Still the results are relatively good during the validation period. That could indicate that there is sense in the developed strategies. Although the strategies pick up identical portfolios during the identifying period, they produce very different portfolios during the validation period. One can conclude that the identification period contains too little information to properly optimize the strategies.

5.3 Optimization Convergence

In this section we test the convergence speed of the genetic algorithm. Two (2) elite sets are used having sizes of 20 and 40. We present the convergence of annual return, Sharpe ratio and maximum drawdown for the best 10 strategies in the elite set during the parameter identification period with respect to the total number of evaluated strategies.

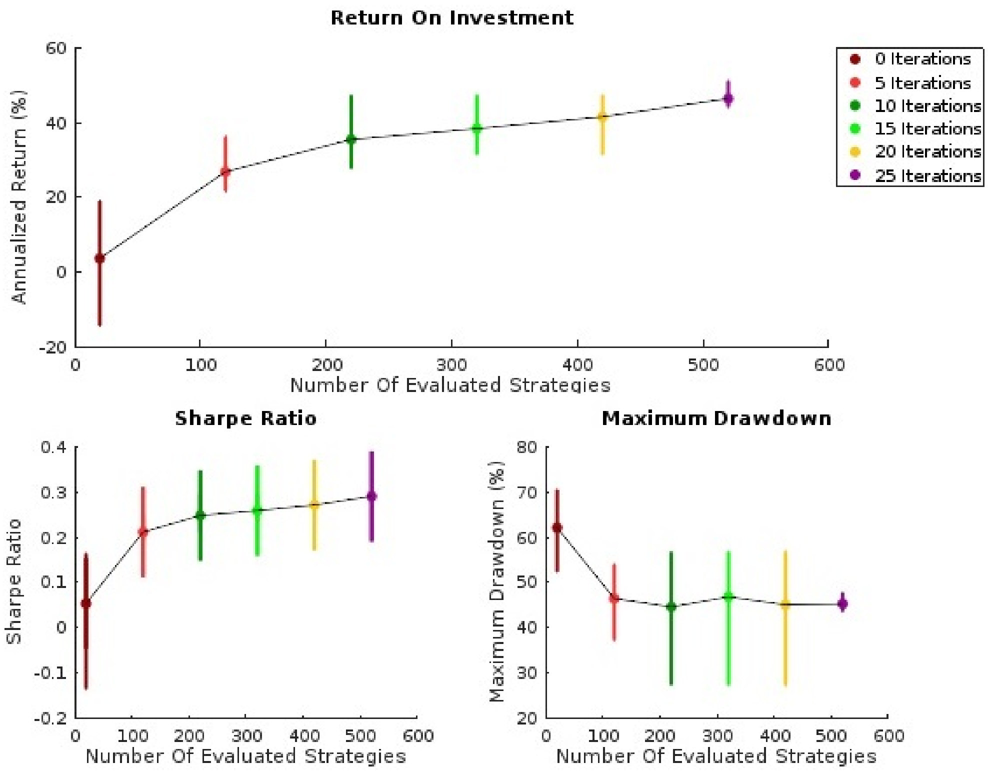

Figure 5.5: Convergence using elite set of 20 strategies

The *-marks in Figures 5.5 and 5.6 present the indicator average values for the elite set. The vertical lines present the range from minimum to maximum values for the indicators that the strategies have in the elite set.

One can note that not every performance indicator in the strategy monotonically improves during the optimization process. Sometimes the average performance of the strategies with respect to one objective may even decrease. This is caused by the multi-objective nature of the objective functional. To

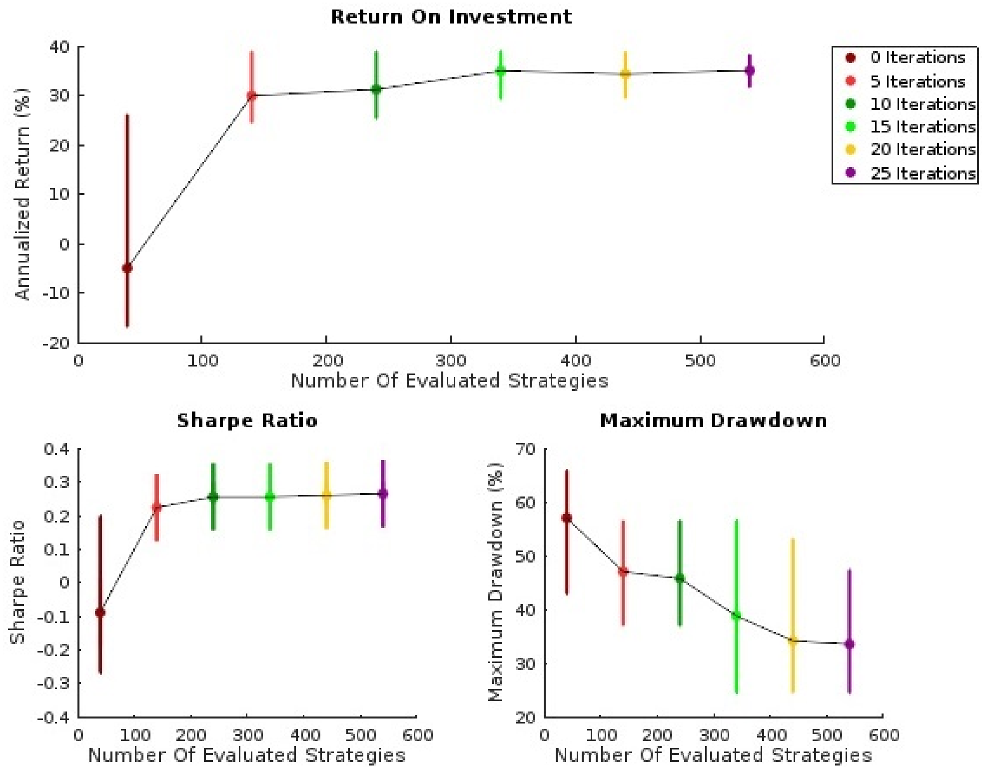

Figure 5.6: Convergence using elite set of 40 strategies

achieve an improvement in one attribute, it may be necessary to permit another to weaken. This phenomenon occurs e.g. in the maximum drawdown of Figure 5.5. The maximum drawdown increases between iterations 10-15, while the return and Sharpe ratio increases. The exchange rate between the attributes depends on the objective functional.

The 20 and 40 size elite sets both converge very similarly. Because of the stochastic nature of the genetic optimization process, it is not possible to draw statistically significant conclusions about the differences.

5.4 Altering the Identification Period

In this section we test how the results differ when the parameter identification period is altered. In the first test (Subsection 5.4.1) a shortened identification period is used, and in the second test (Subsection 5.4.2) the identification period is extended.

5.4.1 Shortened Identification Period

The strategies are medium optimized and the identification period start in May 2009 and ends in December 2011.

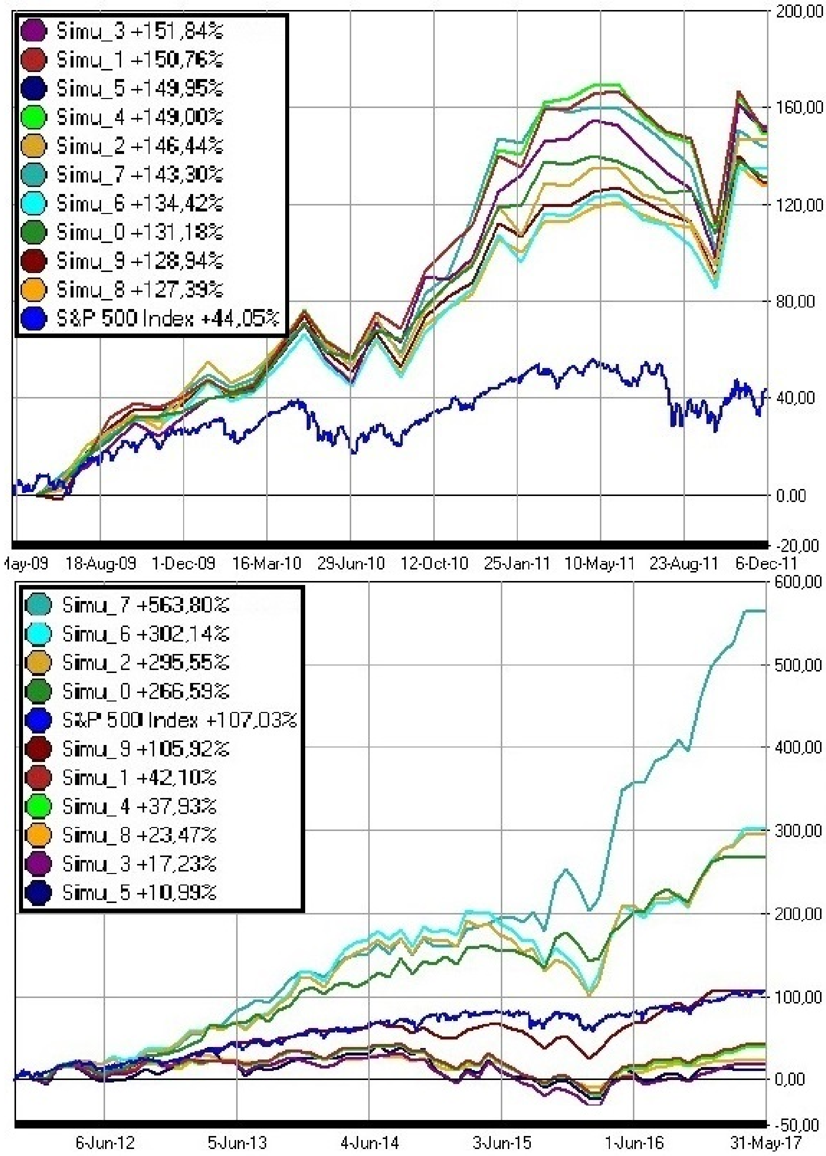

Figure 5.7: Return on investment (%), shortened identification period

According to Figure 5.7 the strategies seem to achieve good performance during the parameter identification period but their performance on the validation period is less encouraging. There seems to be some strategies that outperform the index significantly but others fail.

5.4.2 Extended Identification Period

The strategies of this test are medium optimized and the parameter identification period is changed to start in May 2009 and to end in June 2014.

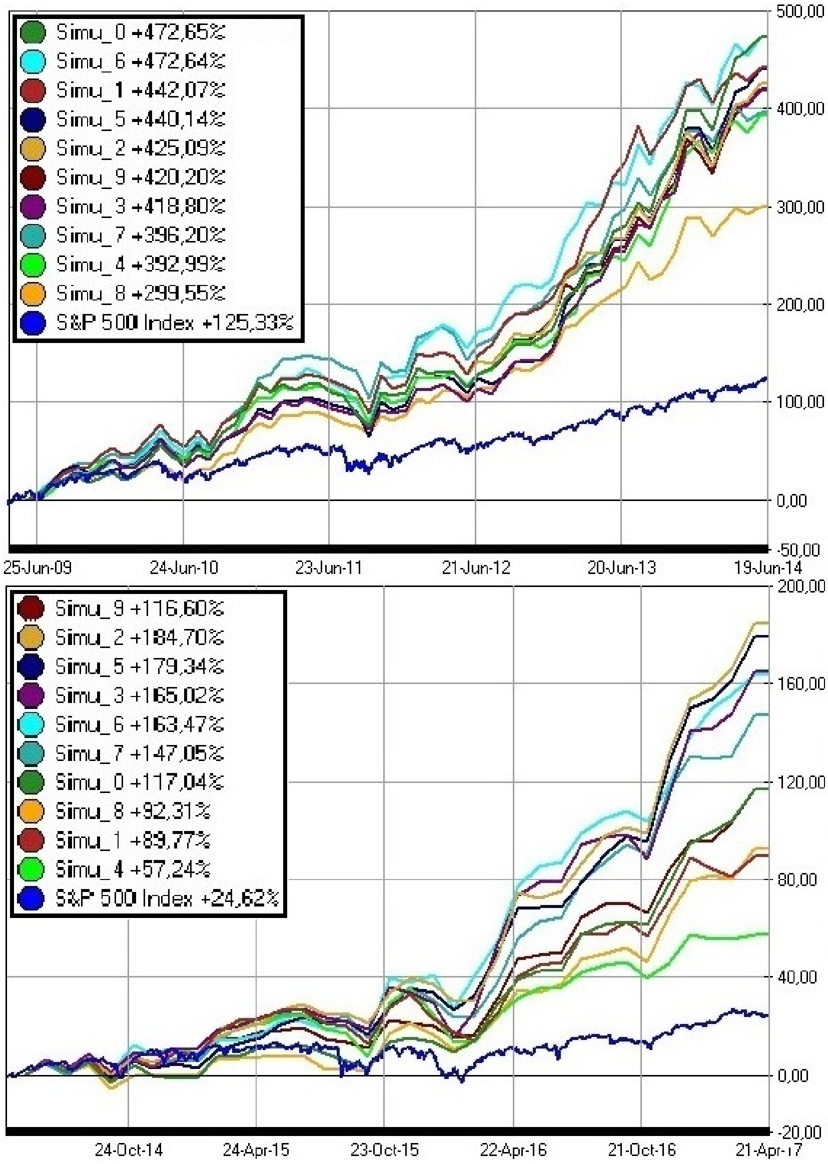

Figure 5.8: Return on investment (%), extended identification period

According to Figure 5.8 the strategies seem to achieve good performance during both the identification and validation periods, the S&P 500 Index is outperformed by every strategy. This could indicate that in the optimization phase more data is needed to avoid fit into sample to occur. The performance during the identification and validation periods is similar indicating that the strategies are not a fit into the sample. The optimized strategies seem to possess stable portfolios.

5.5 Split the Instruments into Subsets

The set of 1378 instruments we have used this far is in this test split randomly to two roughly equal sized sets. The first set is used for the parameter identification, and the second for validation. Both the identification and validation periods span from June 2009 to April 2017. Strategies are medium optimized in this test. This test measures the performance of the genetic optimization method.

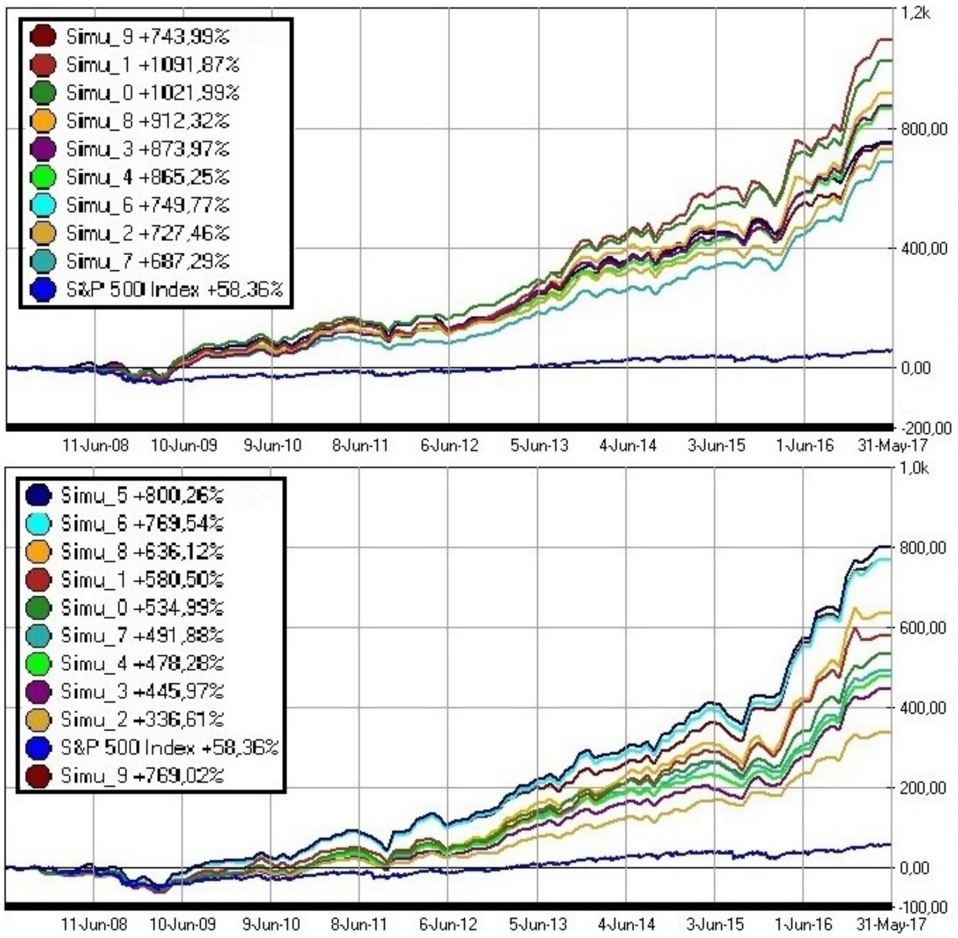

Figure 5.9: Return on investment (%), split

Every strategy in the elite set seems to outperform the S&P 500 Index significantly. The performance is overwhelming with the instruments used in the identification and almost as good with the validation set. The optimization method seems to find very decently working strategies. It is possible that the strategies are a fit into the sample because instruments could have acted similarly due to the identical time periods of the identification and validation sample. One cannot show if there is any real sense behind the strategies. Figure 5.9 reveals only that the used genetic optimization method has good performance and it works very well.

5.6 Sensitivity of the Strategies

The choice of forecasting parameters (decision-making variables) has an effect on the behavior of the strategies. This section tests how sensitive the decision-making rules are to changes.

The set of genes that the strategy consists of have for testing purposes been divided into 2 categories. The first category consists of genes performing the *technical analysis*, while the second group of genes performs *fundamental analysis*. To have a proper sensitivity analysis the method is made partially linear meaning that both these scoring functions have their own contribution weights that can be varied. The structure of the decision-making scoring function used for the study is of the form:

$$S = c_1 \Psi^{\text{score}}_{\text{Technical}} + \underbrace{c_2}_{=1} \Psi^{\text{score}}_{\text{Fundamental}}, \tag{5.1}$$

where $\Psi^{\text{score}}_{\text{Technical}}$ and $\Psi^{\text{score}}_{\text{Fundamental}}$ are the *technical* and the *fundamental* scoring for the instrument and c_1 is the exchange coefficient between the *fundamental analysis* and *technical analysis*.

The focus of the test is to determine whether it is $\Psi^{\text{score}}_{\text{Technical}}$ or $\Psi^{\text{score}}_{\text{Fundamental}}$ that has more contribution to the instrument selection and to portfolio features. It is still possible that the method cannot be divided into two separate parts in the way presented in (5.1) because there can be joint effects between *fundamental analysis* and *technical analysis*.

A medium optimized elite set has been used during the optimization stage of this test. This ensures that both $\Psi^{\text{score}}_{\text{Technical}}$ and $\Psi^{\text{score}}_{\text{Fundamental}}$ are decently converged and functional but are not over fitted to the sample.

The *Technical analysis* is given a larger weight in the first test ($c_1 = 2.00$) of Figure 5.10. *The fundamental analysis* is given a higher weight in the second test ($c_1 = 0.50$) of Figure 5.11. In principle setting $c_1 = 1$ would means a neutral situation where the *fundamental analysis* and the *technical analysis* are weighted in the best way selected by the genetic algorithm.

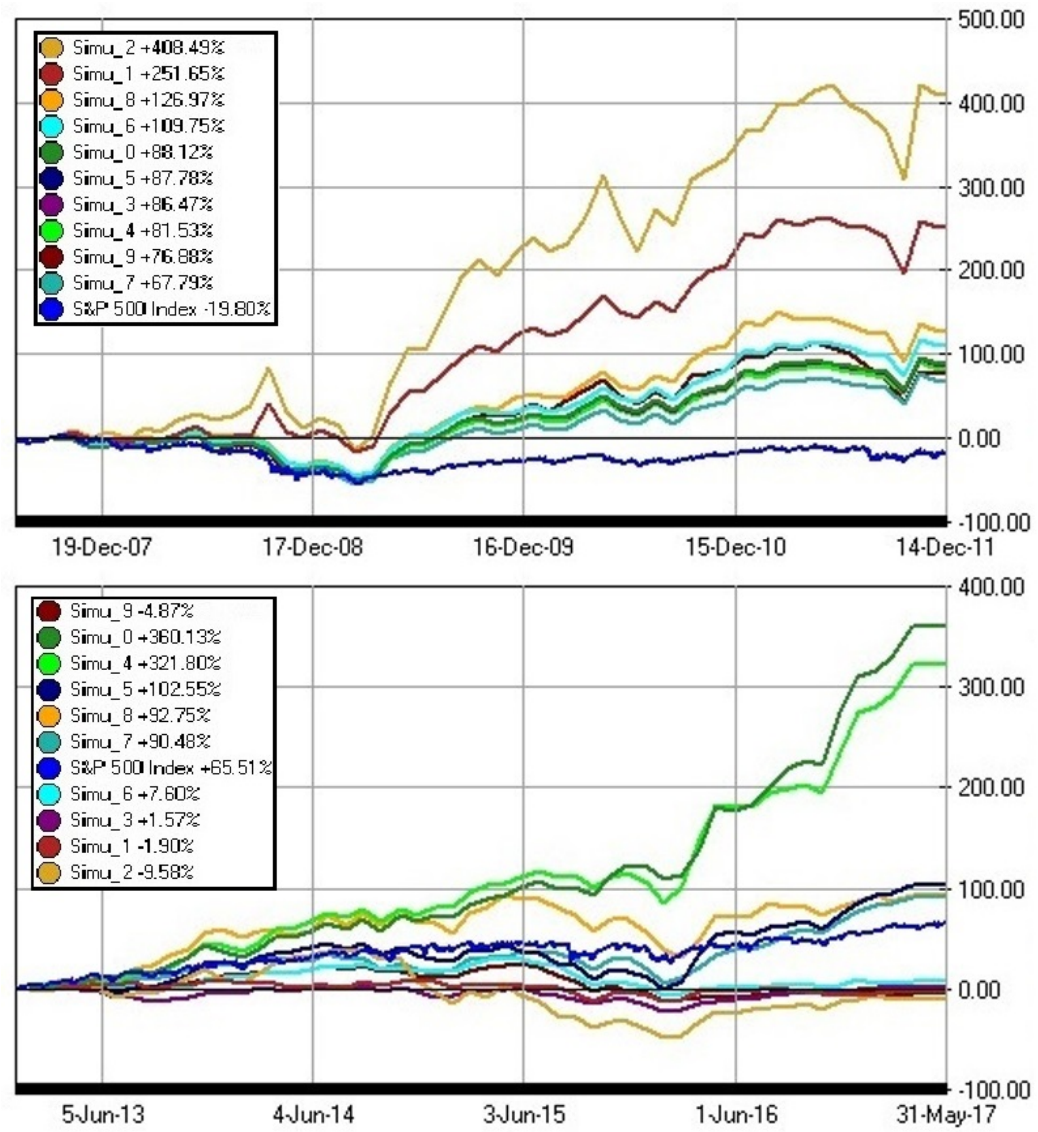

Figure 5.10: Return on investment (%) technical analysis weighted ($c_1 = 2$)

The portfolios of the technically weighted (Figures 5.10) and the fundamentally weighted (Figure 5.11) strategies are unexpectedly almost identical. This means that the value of c_1 does not effect the decision-making. A conclusion can be drawn that the method uses effectively either only *fundamental* or *technical analysis* in the decision-making. Our tests revealed that in this case the method only uses *technical analysis* that was shown by ignoring

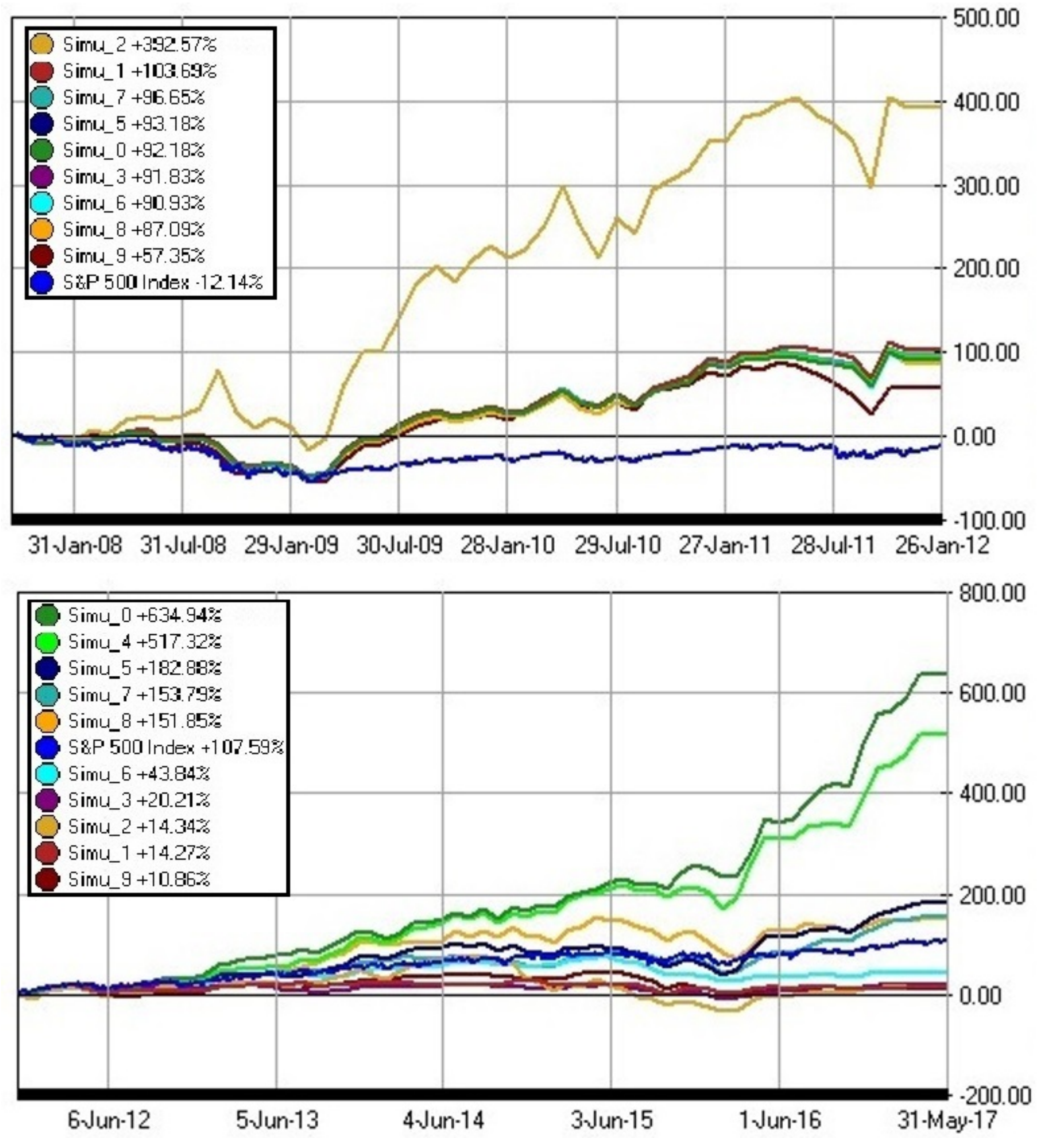

Figure 5.11: Return on investment (%) fundamental analysis weighted ($c_1 = 0.5$)

the *fundamental analysis* by setting $c_1 = 1000$ so that the *technical analysis* becomes the only effective factor.

5.7 Separate Strategies for Technical and Fundamental Analysis

Because of the system nature one must not draw strict conclusions from the sensitivity analysis presented in Section 5.6. The strategies are nonlinear and can differ significantly from strategies where strategy components effect linearly and component-wisely. A nonlinear system can in this sense be more than its components. To illustrate this phenomenon two separate strategies the first based on only *fundamental analysis* and the second based on only *technical analysis* are separately optimized and studied in this section. Total 550 strategies have been evaluated in the optimization stage.

5.7.1 Strategies Based on Fundamental Analysis

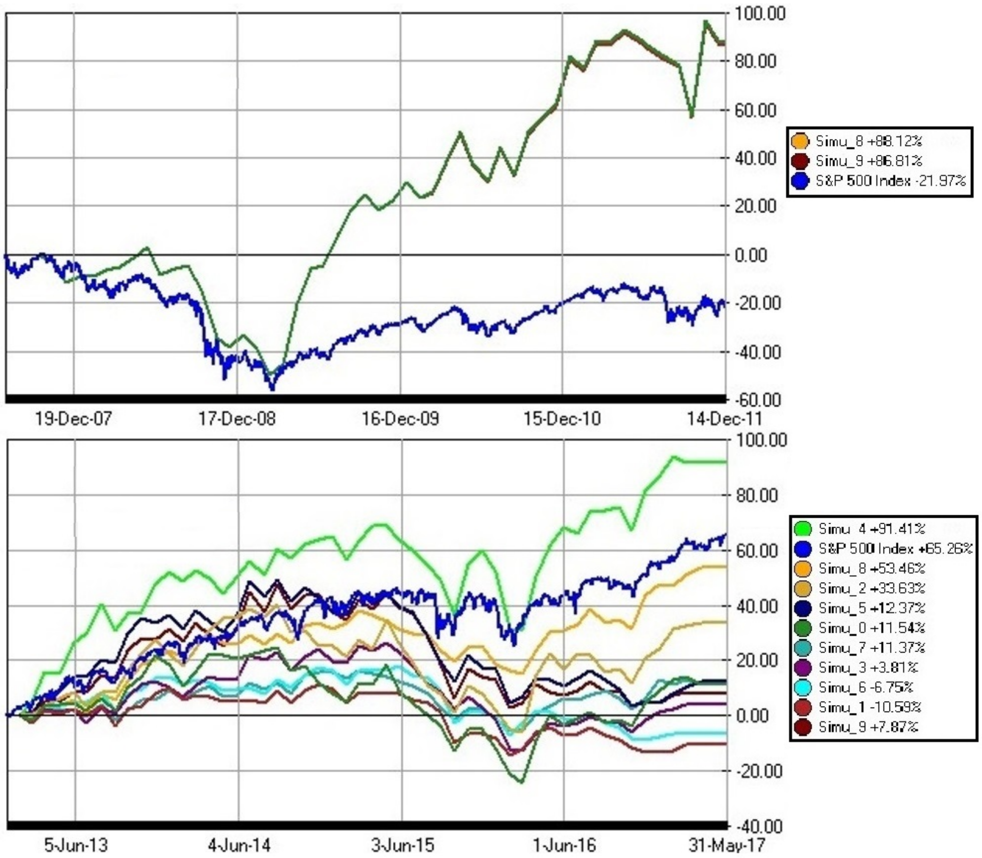

Figure 5.12: Fundamental analysis strategies

From Figure 5.12 one can note that the strategies based only on *fundamental analysis* converge relatively fast. During the training period the optimized strategies have ended up selecting portfolios from two (2) possible choices. This indicates that all the information is used. Interesting is that the strategies still produce very different portfolios during the validation periods. This is a good example from a sample fit. In this case the optimization algorithm works well because it has converged fast during the identification period but the strategies are however undesired fits into the sample.

5.7.2 Strategies Based on Technical Analysis

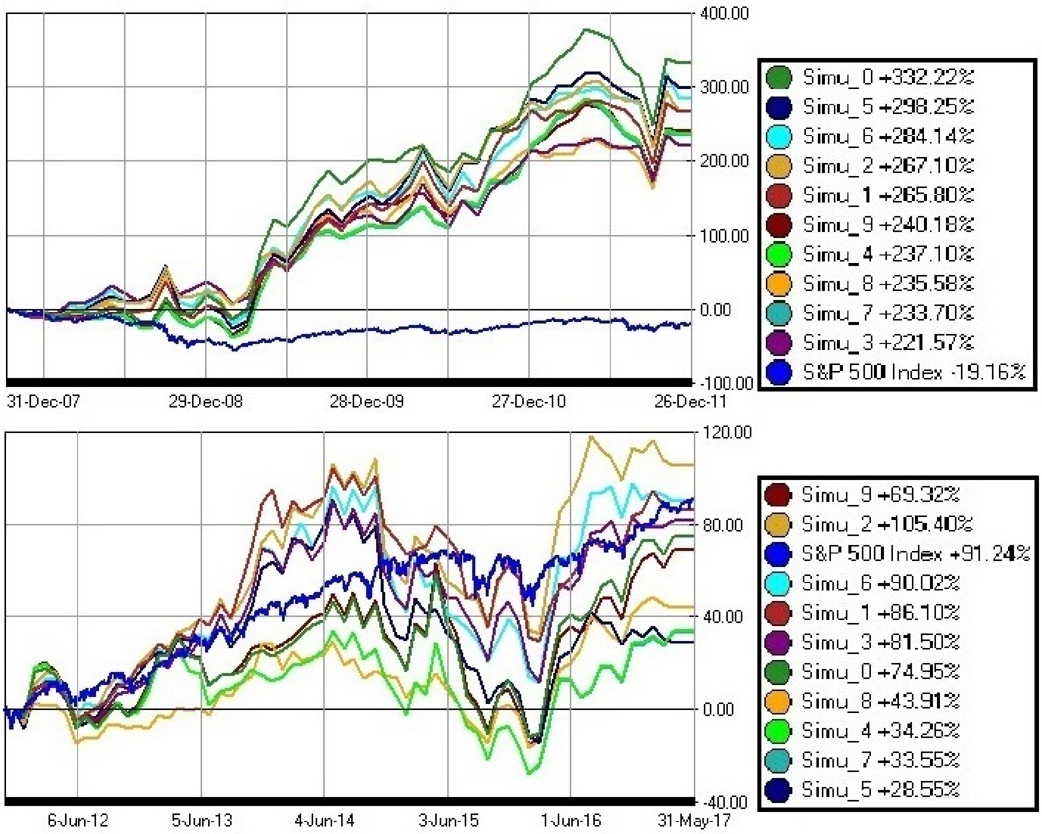

Figure 5.13: Technical analysis strategies

In Figure 5.12 the genetic algorithm has found well working strategies during the training period having diversified portfolios, but the strategies do not perform particularly well during the validation period.

5.8 Conclusions

This chapter discussed the constructed decision-making method and the genetic algorithm performance while using real stock market and economic calibration and validation data.

The strategies and their behavior was tested while varying the amount of optimization. The optimization convergence was studied by measuring how annualized return, Sharpe ratio and maximum draw-down of the strategies converge into some saturation values. Strategy behavior is tested in situations where the period of the parameter identification phase is varied. The sensitivity of the strategies is tested regard to the decision-making parameters that are used to select the instruments. Separate decision-making models are tested, the first based only on *fundamental analysis*, and the second based only on *technical analysis*. At this point the reader is aware of the method behavior and the quality of results it produces.

Chapter 6

Discussion

This chapter studies the used genetic algorithm and the features possessed by the implemented decision-making model. Further improvements for the system are discussed.

6.1 Genetic Optimization Method

The results reveal that the implemented genetic algorithm fits well into finding and optimizing decision-making rules. The instrument selection rules of the strategies stabilize quickly, when the genetic algorithm is applied to optimize them. The converged strategies possess high return and low risk during the parameter identification period indicating that the implemented genetic algorithm performs well.

6.2 Features of the Decision-Making Method

The developed method to form investment strategies is modular, consisting of selection of the forecasting indicators and their parameters, investor preferences function called objective functional and the optimization stage, where the portfolio selection strategies are formed. The system can recognize and pick different kind of well working instrument selection rules (investment strategies) by getting feedback about how satisfying a rule is. It may be hard to tell the explicit idea behind the formed instrument selection rules. The program can create simple working programming code and algorithms on its own and it adjusts the decision-making algorithm, when new information emerges.

It would be more desirable to generate a structure that has more intelligence by granting the structure more freedom regarding to the parameters and functions it can use. While the investment strategy instrument selection in this work was made complicated enough, it turned out that it started to find some algorithms that worked just by chance during the identification period, but had no meaning outside of that sample. The degrees of freedom need to be adjusted to match with the amount of information that the data used for the identification can possess, otherwise it is not guaranteed that the strategy works outside the sample data.

The decision-making rules can, during the optimization stage, take advantage of some factors that do not directly appear in the decision-making indicators. The method does e.g. explicitly try to build a portfolio that allocates money to different sectors. But the optimized strategies end up doing so while preferring strategies having a high Sharpe ratio and a low maximum drawdown. These features are usually maximized by using a good sector allocation strategy. The developed method achieve this goal by optimizing the decision-making to fulfill a proper objective function. Still one must note that the decision-making rules are simple and it is unlikely that they possess very deep intelligence.

It is common that humans are fixated to thought patterns in their decision-making and they do not easily try out abnormal methods and ideas in their decision processes, leading into slow development of new ideas. One advantage of the developed decision-making rule optimizer is that it does not get stuck into thought patterns. Every possible way of thinking is equal for the system regardless of background of the idea. New working ideas have a good chance to show up. A disadvantage of this approach is that it may sometimes lead to absurd strategies working by a good luck, that must be ignored later in the light of the best knowledge and new data.

The performance of the strategies was less perfect outside the sample data, when the sample used for optimization of the decision-making rules was not large enough. This could have resulted from having too many degrees of freedom, enabling an undesirable fit into the identification sample period. An alternative reason could be that the set of forecasting parameters did not possess all the influencing factors, and a better forecast cannot be achieved without increasing the amount and diversity of the indicators. The set of indicators could not be enlarged in the scope of this work, because we did not have time series data for more independent factors available.

6.3 Multi-Stage Selection

The instrument selection decision-making process in the developed system is a one-step procedure. An alternative way would be to implement a multistep procedure, i.e. the strategy preselects for example $N \approx 25$ - 100 instruments from a given set. The next step would be to create the portfolio from these preselected set with another method, e.g. Markowitz portfolio selection method. The advantage would be that with a significantly smaller set N, it will be possible to take explicitly into account the cross correlation or other cross features in the decision-making process. Our set consisted of 1378 instruments and the instrument's cross features was forced to be ignored to make the method simulation and optimization practical.

6.4 Augmenting the Entropy Principle

The decision-making rule could be extended to have more degrees of freedom by creating an entropy penalty factor into the objective functional. Entropy measures the amount of information in the system or in other words measures the system complexity. The lower the entropy is, the better the method is in terms of complexity. Good methods or models work well and can be stored using low amount of information. When using the entropy penalty, one could grant the method a lot more freedom in the building stage and only the best performing sub parts for the final method would be selected. One form of this idea is introduced by Jorma Rissanen.[23]

The entropy penalty is already integrated into e.g. natural selection, where complicated structures are more probable to perish e.g. as a result of thermal motion or ionizing radiation. The complicated structures can survive only if they outperform the simpler ones with a large enough margin.

To generate a method, a decision-making model or an investment strategy that has this property, requires the formulation of the entropy of the system into the objective function.

Chapter 7

Conclusions

7.1 Achieved Objectives

The first objective of this work (i) was to study the type of models, approaches and indicators used in financial decision-making when determining the base and principles in stock investment strategies and forecasting companies future success. The second objective (ii) was to construct and document a structure that produces decision-making rules (investment strategies) that find well performing stock market portfolios to satisfy the investor needs. The third objective (iii) was to implement the structure, optimize it using genetic algorithm, and test its performance with real stock market data, forming investment strategies for stock instrument selection into practical use.

All the objectives of this work (i, ii and iii) were met. Wide spectrum of models and indicators was found and some advantages of them was used, and kept in mind, when the structure of the decision-making rule generator was constructed. The structure was implemented, and it turned out to be successful in generating decently working investment strategies.

7.2 Discussion

The strategies seem often to fit into the sample meaning that they work only with the sample used to optimize the strategies, and does not work outside of this sample with another sample. This happens although the degrees of freedom is comparably low relative to the sample size. Partly this could be caused by some rapidly changing phenomena in the investment world. Still the strategies seem to be well working when they are carefully optimized with sufficient amount of data. It seems that it is important to ensure that the period used to optimize the strategies is long enough. Better investment strategies are obtained when the period is longer.

When the method dynamics is adjusted properly, it adapts rapidly to new phenomena. With as such or small improvements and fixes, the strategies can be very useful and practical in the modern investment world. If the strategies are used to make real investments, it is highly recommended to optimize them with a long period that contains as new information as possible, speaking of this day or yesterday stock prices and other possible information. This ensures that the strategies can adapt to rapidly evolving phenomena appearing in the stock market and take advantage of them.

We used genetic algorithm that imitates the evolution process and this approach turned out to work well in calibrating and forming the decision-making rules. Test results showed a relatively good match. The genetic algorithm turned out to be efficient, when the task was to find working investment strategies using a test data covering 1378 instruments, each having time series histories for many different properties.

It is important to take into account the type and behavior of instruments set used in the model calibration because different instrument sets may possess different behavior and phenomena. In this work strategies are optimized to work in New York Stock Exchange and may not perform so well e.g. in Shanghai Stock Exchange. The model must always be calibrated using same or similar instruments that the model is then applied.

Comparison to other models like Markowitz model was omitted because we did not have the necessary resources to perform the Markowitz analysis to the same test material, and the Markowitz model is very different to our model making the comparison also partly impolitic.

We noticed that there is an entropy formulation for optimal degrees of freedom given to a decision-making rule with respect to the sample size used in the calibration, the minimum described length principle (MDL). We did not study the exact formula for this law and that could be a part of later development.

Bibliography

[1] *Group of Mr J D Rimington*, The Tolerability Of Risk From Nuclear Power Stations. *CB. Health & Safety Executive.*

[2] *Markowitz, H.*, 1952. Portfolio selection. *The journal of finance*, 7(1), pp.77-91.

[3] *Knuth, D.E.*, 1997. *The art of computer programming, 3rd edn. seminumerical algorithms, vol. 2.*

[4] *Duffie, D. and Pan, J.*, 1997. An overview of value at risk. *The Journal of derivatives*, 4(3), pp.7-49.

[5] *Pflug, G.C.*, 2000. Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization* (pp. 272-281). Springer US.

[6] *Yu, L., Wang, S. and Lai, K.K.*, 2008. Portfolio Optimization Using Evolutionary Algorithms. In *Reflexing Interfaces*: The Complex Coevolution of Information Technology Ecosystems (pp. 235-245). IGI Global.

[7] Optinoinnin perusteet, (Principles of Optimization), Luento 11: Rajoitusehdot. Ulkopistemenetelmät ja sisäpistemenetelmät.

[8] *Lai, T.L., Xing, H. and Chen, Z.*, 2011. Mean-variance portfolio optimization when means and covariances are unknown. *The Annals of Applied Statistics*, pp.798-823.

[9] *Scherer, B., 2002.* Portfolio resampling: Review and critique. *Financial Analysts Journal*, 58(6), pp.98-109.

[10] *Sornette, D. and Lapeyre, B.*, Portfolio Optimization and Genetic Algorithms.

[11] *Aparicio, F. and Estrada, J.*, 1997. Empirical Distributions of Stock Returns: Scandinavian Securities Markets, 1990-1995. *Unpublished manuscript*, Carlos III University.

[12] *Anagnostopoulos, K.P. and Mamanis, GG,.* 2011. The mean-variance cardinality constrained portfolio optimization problem: An experimental evaluation of five multiobjective evolutionary algorithms. *Expert Systems with Applications*, 38(11), pp.14208-14217.

[13] *A.E. Eiben, M. Schoenauer b,.* 2002. Evolutionary computing. *Information Processing Letters.*

[14] *Chen, J.S., Hou, J.L., Wu, S.M. and Chang-Chien, Y.W.*, 2009. Constructing investment strategy portfolios by combination genetic algorithms. *Expert Systems with Applications*, 36(2), pp.3824-3828.

[15] *Haddock, J. and Mittenthal, J.*, 1992. Simulation optimization using simulated annealing. *Computers & industrial engineering*, 22(4), pp.387-395.

[16] *Shi, Y. and Eberhart, R.*, 1998, May. A modified particle swarm optimizer. In Evolutionary Computation Proceedings, 1998. *IEEE World Congress on Computational Intelligence.*, on (pp. 69-73)

[17] *Li, X. and Engelbrecht, A.P.*, 2007, July. Particle swarm optimization: an introduction and its recent developments. In Proceedings of the 9th annual conference companion on Genetic and evolutionary computation (pp. 3391-3414). ACM.

[18] *Golmakani, H.R. and Fazel, M.*, 2011. Constrained portfolio selection using particle swarm optimization. *Expert Systems with Applications*, 38(7), pp.8327-8335.

[19] *Soleimani, H., Golmakani, H.R. and Salimi, M.H.*, 2009. Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Systems with Applications*, 36(3), pp.5058-5063.

[20] *Tsai, T.J., Yang, C.B. and Peng, Y.H.*, 2011. Genetic algorithms for the investment of the mutual fund with global trend indicator. *Expert Systems with Applications*, 38(3), pp.1697-1701.

[21] *Chang, Y.H. and Lee, M.S.*, 2017. Incorporating Markov decision process on genetic algorithms to formulate trading strategies for stock markets. *Applied Soft Computing*, 52, pp.1143-1153.

[22] *Mousavi, S., Esfahanipour, A. and Zarandi, M.H.F.*, 2014. A novel approach to dynamic portfolio trading system using multitree genetic programming. *Knowledge-Based Systems*, 66, pp.68-81.

[23] *Rissanen, J.*, 1978. *Modeling by shortest data description. Automatica*, 14(5), pp.465-471.

[24] *Glasserman, P., Heidelberger, P. and Shahabuddin, P.*, 2000. Efficient Monte Carlo methods for value-at-risk. *IBM Thomas J. Watson Research Division.*

Appendix A

Covariance Matrix is Positive Definite

Let $\mathbf{w}$ be the instruments allocation and $\mathbf{\Sigma} = \mathbb{E}\left[(r - \mathbb{E}[r])(r - \mathbb{E}[r])^T\right]$ be the covariance matrix of instrument returns. Now for every portfolio covariance matrix $\mathbf{\Sigma}$ it holds that:

$$\begin{aligned}\mathbf{w}^T\mathbf{\Sigma}\mathbf{w} = \mathbf{w}^T\mathbb{E}\left[(r - \mathbb{E}[r])(r - \mathbb{E}[r])^T\right]\mathbf{w} &= \mathbb{E}\left[\mathbf{w}^T(r - \mathbb{E}[r])(r - \mathbb{E}[r])^T\mathbf{w}\right] \\ &= \mathbb{E}\left[\left\{\mathbf{w}^T(r - \mathbb{E}[r])\right\}\left\{\mathbf{w}^T(r - \mathbb{E}[r])\right\}^T\right] = \mathbb{E}\left[xx^T\right] \geq 0,\end{aligned} \tag{A.1}$$

where $x = \mathbf{w}^T(r - \mathbb{E}[r])$.

If we also assume that the covariance matrix is of the full rank (that is commonly the case) meaning that there are no portfolios with zero variance then it strictly holds that $\mathbf{w}^T\mathbf{\Sigma}\mathbf{w} > 0$.

Appendix B

MACD Formula

$$\beta_i = 1 - \alpha_i \ \forall i \in \{1, 2, 3\}.$$

$$\begin{aligned}
\mathrm{MACD}(\alpha_1, \alpha_2, \alpha_3)[p_t] &= \mathrm{EMA}(\alpha_3)\left[\mathrm{EMA}(\alpha_1)\left[p_t\right] - \mathrm{EMA}(\alpha_2)\left[p_t\right]\right] \\
&= \frac{\sum_{j=0}^{\infty}(1-\alpha_3)^j \left[\frac{\sum_{i=0}^{\infty}(1-\alpha_1)^i p_{t-i-j}}{\sum_{i=0}^{\infty}(1-\alpha_1)^i} - \frac{\sum_{i=0}^{\infty}(1-\alpha_2)^i p_{t-i-j}}{\sum_{i=0}^{\infty}(1-\alpha_2)^i}\right]}{\sum_{j=0}^{\infty}(1-\alpha_3)^j} \\
&= \frac{\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta_3^j\beta_1^i p_{t-i-j}}{\sum_{i=0}^{\infty}\beta_1^i\sum_{j=0}^{\infty}\beta_3^j} - \frac{\sum_{j=0}^{\infty}\sum_{i=0}^{\infty}\beta_3^j\beta_2^i p_{t-i-j}}{\sum_{i=0}^{\infty}(1-\alpha_2)^i\sum_{j=0}^{\infty}(1-\alpha_3)^j} \\
&= \frac{\sum_{n=0}^{\infty}\sum_{k=0}^{n}\beta_3^k\beta_1^{n-k} p_{t-n}}{\sum_{i=0}^{\infty}\beta_1^i\sum_{j=0}^{\infty}\beta_3^j} - \frac{\sum_{n=0}^{\infty}\sum_{k=0}^{n}\beta_3^k\beta_2^{n-k} p_{t-n}}{\sum_{i=0}^{\infty}\beta_2^i\sum_{j=0}^{\infty}\beta_3^j} \\
&= \frac{(1-\beta_3)(1-\beta_1)\beta_1}{\beta_1-\beta_3}\sum_{n=0}^{\infty}\left(1-\beta_3^{n+1}\beta_1^{-n-1}\right)\beta_1^{n+1}p_{t-n} \\
&\quad - \frac{(1-\beta_3)(1-\beta_2)\beta_2}{\beta_2-\beta_3}\sum_{n=0}^{\infty}\left(1-\beta_3^{n+1}\beta_2^{-n-1}\right)\beta_2^{n+1}p_{t-n} \\
&= \frac{(1-\beta_3)(1-\beta_1)\beta_1}{\beta_1-\beta_3}\sum_{n=0}^{\infty}\left(\beta_1^{n+1}-\beta_3^{n+1}\right)p_{t-n} \\
&\quad - \frac{(1-\beta_3)(1-\beta_2)\beta_2}{\beta_2-\beta_3}\sum_{n=0}^{\infty}\left(\beta_2^{n+1}-\beta_3^{n+1}\right)p_{t-n}
\end{aligned}$$

$$= (1-\beta_3)\sum_{n=0}^{\infty}\left[\frac{1-\beta_1}{1-\dfrac{\beta_3}{\beta_1}}\beta_1^{n+1} - \frac{1-\beta_2}{1-\dfrac{\beta_3}{\beta_2}}\beta_2^{n+1} - \left(\frac{1-\beta_1}{1-\dfrac{\beta_3}{\beta_1}} - \frac{1-\beta_2}{1-\dfrac{\beta_3}{\beta_2}}\right)\beta_3^{n+1}\right]p_{t-n}. \tag{B.1}$$